Published in final edited form as:

Appl Nurs Res. 2007 February ; 20(1): 39–43.

Data Quality Strategies in Cohort Studies: Lessons from a Study on Delirium in Nursing Home Elders

Mary J. Dyck, PhD, RN, BC, LNHA,

John A. Hartford Foundation Building Academic Geriatric Capacity, Post-doctoral Fellow and Assistant Professor, Mennonite College of Nursing at Illinois State University, Campus Box 5810, Normal, IL 61790-5810, 309-438-2547 (work), 309-454-9498 (home), 309-438-2288 (fax), mjdyck@ilstu.edu

Kennith Culp, RN, PhD [Associate Professor], and

The University of Iowa College of Nursing, NB 462, Iowa City, IA 52242, 319-335-7060 (work), kenculp@uiowa.edu

Pamela Z. Cacchione, PhD, RN, CSGNP [Associate Professor]

Saint Louis University School of Nursing, 221 N. Grand Blvd., St. Louis, MO 63103, 314-977-8934 (work), cacchipz@slu.edu

Abstract

Data quality has a direct impact on reliability and validity, however, the procedures are usually briefly summarized in the methods section of reports. Sustaining data quality and integrity over time can pose serious challenges and prompted the development of a Data Quality Program based on Donabedian's quality framework. Although many are familiar with the structure, process, and outcome components in health care quality, application to a research project may be unfamiliar. This article summarizes the data quality program for a cohort study of nursing home elders with delirium by providing an "insider's view" of the procedures and protocols followed over several years.

Keywords

quality improvement; data quality; reliability; validity; cohort studies; methodology; nursing homes; delirium

Data Quality Strategies in Cohort Studies: Lessons from a Study on Delirium in Nursing Home Elders

The quality of data collected over a sustained period of time in a longitudinal study impacts on the reliability and validity of measurements. Sustaining data integrity over time poses challenges, especially with staff changes, transcription errors, and historical events. These concepts in a cohort study with elders are not discussed in the literature. Elders may be lost to follow-up due to illness, transfer, and death. This article describes data quality strategies to uphold reliability and validity in a cohort study of nursing home elders.

Supported in part by grants from the National Institute on Aging RO1 AG17939 and National Institute of Nursing Research T32NRO7058 and F31NR07824-01A1.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conceptual Framework

Typically data quality procedures in research projects are based on the assumption that data integrity is a function of reliability and validity. However, the strategies used to maintain reliability and validity over time are rarely discussed.

Standards

Standards are the foundation for quality improvement (QI) programs (Njie & Thomas, 2001). In research, standards are determined by the principle investigator (PI) based on research principles, National Institute of Health (NIH) guidelines, and human subject protections (Morse, Califf, & Sugarman, 2001). Quantitative research standards entail both rigor and control. Rigor involves discipline, attention to detail, and accuracy. Design characteristics for control include sample, setting, and protocols (Burns & Grove, 2001).

Standards, which are often determined by a funding agency, may include budget management, training, study schedules, reports submission, and publications (Burns & Grove, 2001). For example, all research team members need to complete human subjects training. In the past, unethical studies have resulted in federal regulations to protect human subjects. While researchers are responsible to know and protect the rights of human research subjects, the IRB ensures the protection of human subjects (Morse et al., 2001).

Research standards are stringent. Integrity of these standards is maintained through the QI process. Few investigators have studied the effectiveness of QI in research protocols, however, some researchers have provided anecdotal evidence of error detection (Njie & Thomas, 2001).

Quality Improvement

In health care quality QI is a quality management program continuously addresses quality with a focus on processes (Bradley & Thompson, 2000). For this paper, data quality program identifies the quality process used to facilitate quality management programs for research.

While Donabedian (Donabedian, 1986) provides a framework for evaluating quality of care, the same framework may evaluate research quality. Structure identifies the study infrastructure developed in the proposal. Structural components include funding support, instruments, human subjects regulations, and staff. The proposal design is critical to prevent jeopardizing the internal validity. Possible sources of internal invalidity to be controlled through the design include history, maturation, testing, mortality, instrumentation, regression, selection, and interaction of selection (Campbell & Stanley, 1963).

The process dimension includes all aspects of the research process (Njie & Thomas, 2001). Standardized procedures assist in the data quality program by preserving data integrity and protocol adherence. Protocol violations of 65.8% were found in one survey (Rankin & Esteves, 1997) suggesting that protocol adherence is a huge research challenge.

In research, outcome is defined as data completeness and accuracy. Results support the analyses making them defensible and generalizable, thus maintaining external validity. If many persons decline to participate in a study, the sample may limit external validity because results may not be generalizable to the full population. Persons with dementia may have paranoia causing them to be skeptical of strangers. As a result, they may refuse to participate. Settings may limit generalizability by those who agree or do not agree to participate (Burns & Grove, 2001). This research team found that rural facilities often did not understand research studies.

Exemplar: Data Quality Program For A Delirium Study Of Nursing Home Elders

This cohort study in rural nursing homes was designed to identify the epidemiology of delirium in nursing home residents (Culp et al., 2004). Research staff were involved in data collection in a 14-county area, but were in a facility for only six weeks. For the first two weeks, Research Assistants (RAs) collected baseline data. Cognitive testing was completed three times a day for the first 14 days after baseline. For the remaining 14 days, cognitive testing occurred daily except weekends. After six weeks, data were sent to the central research office and the RAs moved to the next site to repeat the process.

This study defined a cohort as persons who have experienced surveillance over a specific time period with data collection occurring over 2 1/2 years. The purpose of the Data Quality Program was to ensure accuracy of cognitive, behavioral, and physiological ratings and data entry under the protocols described elsewhere (Culp et al., 2004). The PI set the data quality standards based on physiological plausible measurements and established the structure for the Data Quality Program, including a nurse Data Quality Coordinator (DQC) and data entry clerks.

Structure

In this study, structural components included the proposal, research sites, instruments, laboratory facilities, and staff. The relationship of these components to the Data Quality Program is described below and in Figure 1.

The proposal defined structural elements including the design, subject eligibility, criteria, sample size, instruments, and location. A two-stage cluster sampling technique was used based on standard epidemiological research design (Kleinbaum, Kupper, & Morganstern, 1982). The intensive surveillance in Days 1 through 14 was based on previous research (Inouye & Charpentier, 1996) that identified no new cases after six days. The use of a standard design and rationales based on the state of the science worked to prevent threats to internal validity.

Data Quality in Sampling

A two stage cluster sampling technique was used to select nursing homes and residents. A final sample of 300 nursing home residents was desired for this study, which required a sample of 360 due to attrition. Initially, random sampling of counties in the 14-county catchment area was based on a list of all facilities in the geographic region. Equal opportunity to respond to the study invitation was provided to all 45 facilities. In the second stage of the cluster sampling, facilities were randomly selected based on size. After facility selection residents were randomly selected based on a comprehensive list of facility residents. Nursing home residents are vulnerable to human rights violations due to physical and psychological problems, including dementia (Njie & Thomas, 2001). Since dementia is a risk factor for delirium, residents were at particular risk. Protocols included random sampling and use of assent for residents' unable to give consent followed by informed consent from a responsible party. Southeast rural Iowa with 14 counties and 45 nursing homes provided the setting. Fourteen nursing homes from the catchment area participated in this study maintaining selection criteria and internal validity.

Data Quality with Instruments

Instrumentation can have a major impact on internal validity. In this study, instrumentation included assessment tools, bioelectrical impedance analysis (BIA) equipment, and weight scales. Assessment tool psychometrics were evaluated during proposal development. BIA use with elders was evaluated at an NIH-sponsored conference with the recommendations used for this study. Facility scales were used for weights, since moving and recalibrating a scale was difficult. Residents were weighed at the same time of day based on medications and individual characteristics.

Data Quality with Staff Training

Each staff position had different responsibilities in the Data Quality Program. The RA positions required experience as a Registered Nurse in long-term care. Nursing home experience assisted in understanding facility organization, functioning, and interacting daily with residents and staff. Accurate and timely data collection within protocols was the responsibility of the RAs. The key to this Data Quality Program was the DQC, who designed, implemented, and managed the program through the 2 ½ years of the study.

Challenges to Data Quality

Challenges to structural aspects of the study were facilities, equipment changes in the contracted laboratory that performed serum measures, and staff turnover. One facility was a significant distance from the office and the RAs resulting in expenses exceeding budget allocations. One site chose to wait until later in the study due to administrative turnover. The contracted laboratory changed its equipment causing changes in normal values. RA turnover resulted in additional recruitment, hiring, and orientation. These problems could have had negative results on both reliability and validity.

Process

The process aspect impacted all of the study protocols (Njie & Thomas, 2001) because of the need to have standardized procedures to protect data integrity and to ensure protocol adherence. Three procedures were critical to the Data Quality Program – competency assessment, inter-rater reliability assessment, and data completeness audits. While competency and inter-rater reliability assessments have similar actions, they served different purposes. Features of both assessments are given in Table 1.

Competency Assessment

RA competency assessment began during orientation. The DQC coordinated orientation for the RAs, which included content on delirium, differences between delirium and dementia, protocols, assessments, instruments, human subjects, subject recruitment, and blood-borne pathogens. A written competency test was used after the orientation. While the written competency test did not require a minimum percentage, an indication of each RA's knowledge was noted. This orientation was crucial to preventing threats to internal and external validity.

Ongoing competency assessment for cognitive testing occurred at least quarterly. The DQC performed this assessment to determine protocol adherence, since an RA who routinely performed the protocol in the field was less likely to identify subtle changes in protocol adherence by colleagues. Subtle changes in protocol administration could have compromised the accuracy of the measurements, or reliability. Having the DQC administer quarterly competency assessments assisted in maintaining external validity. Discussions of measurement issues, including administration technique, scoring, and interpretation, arose from these competency observations with both the Project Director and PI. These leadership team meetings brought about solutions and led to uniform data collection among the RAs. This process contributed to external validity.

The effort to maintain ongoing competency assessment also included in-service education, which occurred twice a year in the research office. These retreat days encouraged RAs to discuss issues as they were not scheduled for data collection on the same days in the field. Regular meetings of the PI, Co-Investigators, Project Director, and DQC provided further opportunities for discussion of measurement issues. Other elements useful for maintaining competency were: 1) videotapes in the field of subjects with delirium that were reviewed by the PI and external expert consultants, 2) baseline and weekly progress reports faxed to the PI,

and 3) immediate consultation with the PI or project director through the use of a cellular phone. All of these procedures provided support for the RAs in knowing that they were not alone in the field.

Inter-rater Reliability Assessment

Inter-rater reliability determined consistency between RAs on cognitive testing and was completed one time per facility or eight times per year. If percent agreement fell below 90%, then remediation occurred. In the case of that particular observation, it was excluded from the dataset.

Monitoring clinical procedures consistently are important to the accuracy of measurements. For example, weights would be less accurate without vigilance in using different scales or delirium could be missed without observation of clinical changes in the mental status of subjects. An example of this is the use of the Mini-Mental State Exam (MMSE) with visually impaired elders who had difficulty copying pentagons and writing sentences. The team decided to use a thick black marker to outline the pentagons as this assisted in visualizing figures.

Data Completeness Audits

The data for each subject were evaluated in the research office for completeness and accuracy. Occasionally, nursing judgment was required in the evaluation process, so an RN, usually the DQC, performed a chart audit. Common errors included math miscalculations and an occasional blank on a form. If possible, data were corrected from recalling field observations. Since the RAs were assigned to a small group of subjects and were in the residents' rooms regularly for 6 weeks, it was often possible to correct data. Occasionally, there were missing data as some subjects refused to respond. Chart audit errors were summarized by the DQC and given to the PI, Project Director, and each RA. The RAs' copies included all errors, but only their own errors were referenced without the rest of the RA staff knowing their identity. Only the leadership team knew the error rate across employees. Each RA knew the global concerns regarding errors and how their own error rate contributed to the summary report. This ongoing feedback assisted the RAs in being vigilant with data quality.

Cognitive testing in this study provided an example of how each of the various aspects of the Data Quality Program were implemented. Through competency assessment RAs became sensitive to protocol adherence and were willing to precisely implement study protocols. With inter-rater reliability assessment, measurement issues were defined, discussed, and standards determined for this study. While chart audits determined math errors, they also revealed assessment patterns. For example, previous delirium research showed an increase in delirium episodes in the evening. With documentation of the times of cognitive assessments, chart audits provided data to assist the PI in setting cognitive assessment times and evaluating protocol adherence. When a pattern of "morning" data collection points was discovered, the PI intervened and directed staff to use more "evening" scores.

Once the chart audits were completed, the data were entered into the database. While data entry may occur using single entry, double data entry, or contracted services, the PI chose to use single entry with online audits by a second data clerk. The choice of method varies depending on available resources, the PI's experience with data management tasks, and the importance of ongoing data entry. Single entry of data was chosen for this study as the PI decided that some data required RN assessment and could not be delegated to a data management department without a nurse. A chart audit by an RN prior to data entry and a database audit by an RN following data entry provided consistency in the nursing interpretation of this data. If there were erroneous measurements, these were corrected by the RN based on the written chart data..

The role of the DQC was critical to the process component of the study. While the DQC designed and implemented the Data Quality Program, ongoing management was vital to the reliability and validity of the study. Competency and inter-rater reliability assessments needed to be completed, as well as the chart and database audits. Without attention to detail and ongoing evaluation, appropriate outcomes would not have been achieved.

Outcome

The outcome of data quality, defined as completeness and accuracy, contributed to the validity of the study. While evaluation of data was completed using descriptive statistics, the essential keys to the outcome were the structures and procedures provided for this study. Analyses can be defended and results are generalizable indicating external validity.

Conclusion

Since data quality impacts directly on the reliability and validity of a study, this Data Quality Program was developed to ensure the quality of the data. Based on Donabedian's (Donabedian, 1986) quality model, the Data Quality Program reflects structure, process, and outcome concepts with implications for reliability and validity as part of each concept. Threats to internal validity are most common in the structural component as these issues relate to the design, sampling methods, and instrumentation. External validity threats occur in the process concept as procedures and protocols affect the ability to generalize results. The outcome concept of data quality is critical to the validity of the study. Once the standards, structures, and procedures for the study are in place, these reliability and validity concepts might not be discussed. They are, however, the foundation for a high quality research study.

References

- Bradley, MG.; Thompson, NR. Quality Management Integration in Long-Term Care: Guidelines for Excellence. Baltimore, Maryland: Health Professions Press; 2000.
- Burns, N.; Grove, SK. The Practice of Nursing Research: Conduct, Critique, and Utilization. 4. Philadelphia: W. B. Saunders Company; 2001.
- Campbell, DT.; Stanley, JC. Experimental and Quasi-experimental Designs for Research. Boston: Houghton Mifflin Company; 1963.
- Culp K, Wakefield B, Dyck MJ, Cacchione PZ, Decrane S, Decker S. Bioimpedance analysis and other hydration parameters as risk factors for delirium in rural nursing home residents. *Journals of Gerontology: Medical Sciences* 2004;59A(8):813–817.
- Donabedian A. Criteria and standards for quality assessment and monitoring. *QRB: Quality Review Bulletin* 1986;12(3):99–108. [PubMed: 3085044]
- Inouye SK, Charpentier PA. Precipitating factors for delirium in hospitalized elderly persons. Predictive model and interrelationship with baseline vulnerability. *JAMA* 1996;275(11):852–857. [PubMed: 8596223]
- Kleinbaum, D.; Kupper, L.; Morganstern, H. Epidemiological Research: Principles and Quantitative Methods. New York: Van Nostrand Reinhold; 1982.
- Morse MA, Califf RM, Sugarman J. Monitoring and ensuring safety during clinical research. *JAMA* 2001;285(9):1201–1205. [PubMed: 11231751]
- Njie VPS, Thomas AC. Quality issues in clinical research and the implications on health policy (QICRHP). *Journal of Professional Nursing* 2001;17(5):233–242. [PubMed: 11559879]
- Rankin M, Esteves MD. Perceptions of scientific misconduct in nursing. *Nursing Research* 1997;46(5): 270–276. [PubMed: 9316599]

Figure 1.
Data Quality Model

Table 1
Comparison of Competency and Inter-rater Reliability Assessments

Function	Competency Assessment	Inter-rater Reliability Assessment
Purpose	To assess ability of RAs to complete testing according to study protocols	To assess ability of RAs to measure consistently between examiners
Timing	Quarterly	One time per facility, i.e. 8 times per year
Evaluator	Data Quality Coordinator	RA
Tests	Two cognitive tests selected by Data Quality Coordinator; tests varied over the year	Two cognitive tests selected by Data Quality Coordinator; tests varied over the year
Procedure	Data Quality Coordinator observed each RA administering the test and independently scoring the subject's responses	An RA was selected to observe another RA administering the test and independently scoring the subject's responses
Scoring	Percent agreement	Percent agreement
Evaluation	Percent agreement must be 90% or higher	Percent agreement must be 90% or higher
Measurement Issue	External Validity	Inter-rater Reliability