Epidemiologic Reviews
Vol. 20, No. 1

Printed in U.S.A.

Quality Assurance and Quality Control in Longitudinal Studies

Coralyn W. Whitney, Bonnie K. Lind, and Patricia W. Wahl

INTRODUCTION

Many elements of quality control for cohort studies are similar to those for other types of studies, i.e., standardization of protocol (study procedures), good communication among all study staff, clear expectations of requirements, and monitoring to ensure that requirements are met. However, cohort studies have the added dimension of longitudinal data collection, which adds issues related to drift over time, changes in equipment (including both degradation due to wear and upgrades due to new technology), and staff turnover. If the study uses central laboratories or reading centers to process data, these sites are particularly subject to quality issues related to changes over time. Cohort studies involving multiple centers have extra issues related to comparability of data collected at different locations. This presentation will provide a general overview of quality control elements common to many types of medical research studies, but primarily focuses on issues related to cohort studies. Emphasis will be placed on multicenter cohort studies, as they face additional quality control issues compared with single-center cohort studies.

Much of the literature on quality assurance and quality control has arisen in the context of clinical trials, focusing on maximizing the quality of the data through standardized study-wide and local protocols, training of study personnel, and data management systems (1–9). Quality assurance and control procedures are generally described in the context of a specific study, such as the Multiple Risk Factor Intervention Trial (10–12), the Hypertension Primary Prevention Trial (3, 13), or the Optic Neuritis Treatment Trial (14, 15), with some articles providing general guidelines (4, 16–18). There is less literature in the context of nonclinical trial studies, although some study-specific approaches (8, 19–22) and general quality assurance/quality control guidelines have been addressed (5, 23–25). The literature specifically addressing quality control issues in longitudinal studies is limited (10, 16, 26).

Clinical trials often have a large number of sites, each collecting a small amount of data on a few individual patients, while cohort studies generally involve fewer sites collecting a much broader set of data. This facet of cohort studies can lead to unique quality control challenges due to the sheer bulk of the data collected (2, 10, 27, 28). Therefore, quality control literature based on clinical trials is often not relevant to cohort studies. By having fewer sites, the cohort model is more conducive to the development of long-term relationships between central quality control supervisors and site personnel, and to the supervision of individual site performance. However, it also means that the consequences of a site with quality control problems (e.g., failure to follow protocol, sloppy data handling procedures) can be much more grave for the study as a whole.

Quality control is one of the most important aspects of any study, as the integrity of the conclusions drawn by a study are in large part determined by the quality of the data collected. Data of poor quality, containing a great deal of random noise, decrease the power of a study and can cause a type II error. An even worse result is the collection of data that are biased due to faulty instruments or errors in implementing the protocol, leading to an incorrect report of relations (type I error). Many aspects of data collection can impact the quality of the data, including completeness and clarity of questionnaires, the interviewer's delivery, the accuracy of mechanical instruments, and technicians' measurement techniques (16, 19, 20, 23). The validity of the study depends on the interviewers and technicians from all centers consistently applying study protocol (3, 16). Other errors can be introduced into the data after the data are collected, during transcription, at data entry, and data manipulation for analysis (1, 2, 29). Minimizing all of these potential sources of error is of paramount importance in the planning and implementation of any study. This presentation will suggest ways of minimizing these potential sources of error.

There are actually two basic components to quality

Received for publication June 26, 1997, and accepted for publication May 8, 1998.
From the Department of Biostatistics, School of Public Health, University of Washington, Seattle, WA.
Reprint requests to Bonnie K. Lind, Department of Biostatistics, Box 35–8223, University of Washington, Seattle, WA 98195.

control: quality assurance and quality control. Quality assurance consists of those activities that take place *prior* to data collection while quality control consists of those activities that take place *during* and *after* data collection to identify and correct any errors or discrepancies in the data that have been collected (table 1). Each of these components will be discussed separately.

QUALITY ASSURANCE

Overview

As stated above, quality assurance consists of those activities undertaken *prior* to data collection to ensure that the data are of the highest possible quality at the time they are collected. These activities include the development of the study protocol, development of the data entry and data management systems, training and certification of data collection personnel, and testing of data collection procedures.

Development of the study protocol

The major component of quality assurance is the development of the study protocol and the creation of manuals documenting the protocol. Aspects of protocol development that are common to most study designs are summarized below (table 1).

Design of data collection instruments, including content, format, and step-by-step instructions for completion of the instrument. Procedures should be designed to ensure that the data being produced are in fact reliable, valid, on the appropriate scale, and do not reflect bias of the instrument or bias that may arise in subgroups of the population being studied (13, 14, 20, 23, 26, 30). For example, older participants tend to give responses that are "socially desirable" rather than strictly accurate (19). Limitations of various types of instruments should be considered during development so that potential problems can be identified and mitigated to the extent possible. For example, issues in-

TABLE 1. Summary of quality assurance and quality control activities

Quality assurance activities	Quality control activities
Development of protocol and recruitment	
Protocol development and design	Checking for participant eligibility
Questionnaires and data collection instruments	
Design	Data cleanup
Pilot testing	Local and central evaluation
Coding and editing procedures	Problems with self-report data
Trouble shooting	
Missing values	
Updating	
Interviewer and technician protocol	
Development/pilot testing	For each interviewer and technician:
Training of interviewers	Observe application of protocol
Training of technicians	Monitor certification maintenance
Certification procedures	Completeness of data
Recertification procedures	Retraining when needed
	Regular feedback on quality control
	Training of new personnel
Data entry packages	
Development/pilot testing	Updating systems
Training of personnel	Double entry
Study wide	
Development of local quality assurance/ quality control	Monthly quality control reports
	Reliability/reproducibility studies
	Site visits
Use and maintenance of study equipment	
Develop and pilot equipment use	Routine maintenance of equipment
Training of personnel	Trouble shooting
Maintenance and schedules	Quality of equipment data

herent in self-administered questionnaires have been described in the literature, such as the lack of reliability of self-reported data and the potential for questions to be misunderstood (23). Adequate testing of self-administered questionnaires on volunteers similar to the anticipated cohort are crucial for identifying any problematic questions before the study begins. Interviewer-administered questionnaires also have potential problems; for example, participants may respond differently to interviewers of different age, gender, or ethnic background. In addition, it is vital that the protocol outlines how interviewers are to respond when participants ask for clarification of a question (i.e., is the interviewer allowed to rephrase the question in his/her own words, and if so, how does the study ascertain that interviewers are rephrasing a question in comparable ways). During planning, investigators should also consider any problems that may arise with regard to cohort members who are illiterate or do not speak English.

Cohort studies typically involve many interviewers, creating the potential for differences between interviewers or centers and for differences over time (due either to drift or staff turnover). This makes the establishment of a documented, standardized interviewing protocol of paramount importance (2, 8, 10–12, 31). In addition, investigators must decide during the planning phase whether questionnaires will be allowed to change over time and how the study will ascertain that responses given over time are comparable.

All aspects of the protocol must be documented in a manual of operations (2). This document should be viewed as the official study reference document for data collection staff and, as such, should contain all details of data collection procedures. Once the manual is written, it should be reviewed so that any sections that are ambiguous or subject to misinterpretation can be clarified (32). The manual should help to ascertain that data collection is consistent across field centers and over time. Many good manuals of operations exist in ongoing multicenter cohort studies, with examples usually available from funding sources (33–35).

Selection of standard equipment for all measured data. All sites should begin the study with identical equipment to minimize data variability due to collection by different equipment. A protocol should be established to maintain and recalibrate equipment over time, and forms should be developed to document that these activities have occurred. Ways of identifying and replacing equipment that is worn out or failing must be agreed upon, and procedures for dealing with the development of new technologies must be discussed.

Development of procedures for reviewing and updating the protocol as needed, and for communicating changes to all study personnel. Once data collection actually begins, minor adjustments to the protocol are usually needed. Oversight of this function should be assigned to a group of investigators, and a protocol should be developed for implementing and communicating such changes to all study personnel (see the subsection on "Communication," below).

Development of procedures for obtaining and maintaining certification to perform study procedures, and procedures for monitoring that requirements are met. An integral component of quality assurance is the training and certification of study personnel in accordance with study protocol (2, 3, 10–14, 16, 32). Procedures must be established for training and certifying data collection personnel prior to the start of the study, as well as for training and certifying new staff hired throughout the course of the study. In addition, requirements for maintaining certification must be set (both in terms of number of procedures performed and the quality of the performance), and procedures must be developed for recertification of any technicians who fail to maintain certification.

Hiring individuals to serve as field center technicians is also a crucial part of quality assurance. The personalities and skills of these individuals will have a direct bearing on the final quality of study data. Some sites have found that data quality is better if technicians are hired who have little or no medical background. A person who has been performing procedures such as blood pressures in a clinical setting for many years may be more difficult to retrain to follow a standard protocol than a person who has never performed blood pressures.

Communication. Communication is a key feature of successful cohort quality control systems (2), and part of the protocol development process should be devoted to setting up the structure of the communication system in the study (figure 1). Many large studies use a committee structure to ensure that tasks are accomplished and problems addressed, and in multicenter studies, the coordinating center plays the key role of facilitating communication between committees and field centers. For example, an operations committee may be formed to address questions that arise about implementation of the protocol, to ensure that the protocol is followed at all sites, and to address any problems that arise. A quality control committee may be formed to monitor the actual data quality and to address any problems that arise. For such committees to be effective, communication routes must be established. These communications often take the form of quality control reports, typically produced by the coordinating center, which are distributed to principal investigators. It is equally important that proce-

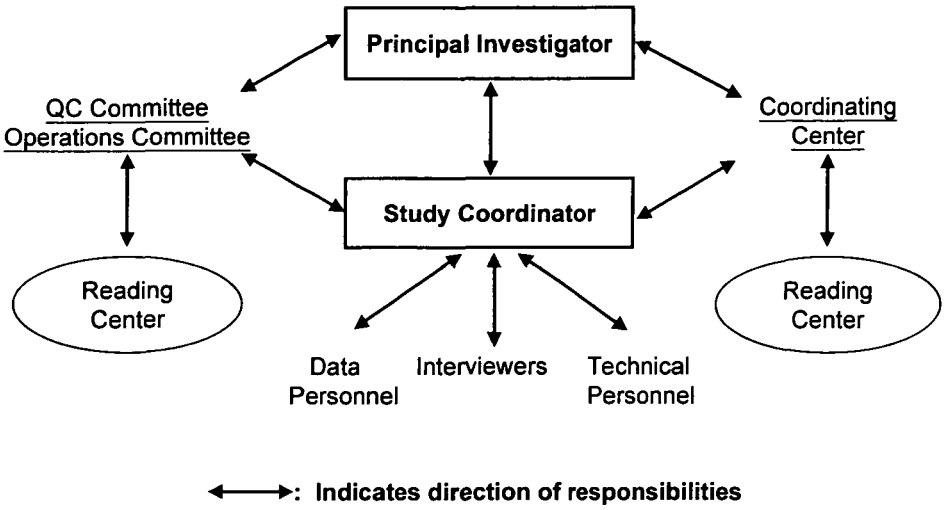

FIGURE 1. Organization and communication among study personnel.

dures be developed for responding to quality control reports, including identification of the person who is to respond to any problems identified in a quality control report, what steps are to be taken, and what the deadline is for response (table 2). During the development phase, the study leadership may also want to delegate responsibility for quality control at each center to a specific individual designated as the quality control supervisor, who could be an investigator, the study coordinator, or another staff member.

Development of data editing, transcription, and entry procedures

Collection of accurate data is only the first step. It is equally important that errors are not introduced in the process of converting the data to electronic format. For some forms it may be desirable to have the form reviewed and edited before data entry. The purpose of this review would be to ascertain that the form was complete, that skip patterns were followed, and that any data values that seemed inconsistent or looked like possible errors were checked before being entered.

Once the data are ready to be converted to electronic format, several options are available. In the past, most studies have relied upon key-entry of data into the computer. Software packages exist that allow data entry to be customized to limit the range of entered data, to check for internal consistency, and to catch errors in key fields such as participant identification numbers. Many studies use double-entry verification to catch and correct any data entry errors from the original entry (16, 36–38). While this system may seem cumbersome, it is more efficient than trying to identify and correct errors at a later point in time (38).

New technologies are continually being developed which offer more efficient ways of converting data to electronic format. Some studies have interviewers enter data directly into a computer without first writing the data onto a paper form (21). In this type of system, it is important that the data entry software include as many ways of checking data accuracy at entry as possible, as any errors identified later cannot be compared with a paper form for correction. Some of the newest technologies being used involve scanning paper forms directly into the computer, or faxing forms to a central data center where the fax machine serves as a scanner to convert the data to electronic format. These systems are also very efficient, but it is important that scanned files be visually reviewed for accuracy.

Training and certification

Once all protocols have been developed and documented, the next step is to train study personnel to implement the procedures. Training and certification activities result in standardization and can lead to reduced costs over time (16). Such training is key, as the interviewer's or technician's perceived value of the data being collected can have a direct impact on the care taken in following the protocol and the likelihood that discrepancies are introduced (27, 32). In a multicenter study, a central training session is often the best way to ensure that all personnel are trained in a standardized manner, providing the added benefit of instilling in local personnel the wider scope of the study and the importance of their work.

Testing of procedures

The final step in the quality assurance process is to test all procedures that have been developed. This is

TABLE 2. Quality control and recruitment reports—responsibilities and distribution

Report	Purpose	Frequency/ distribution*	Committee responsible	Distribution	Response required	
					From	Schedule
Recruitment and data completeness report	Tabulates recruitment and study data received at coordinating center by center	Weekly (F)	Operations committee	Principal investigators, program office, study coordinators	Principal investigators	As needed
Failed studies	Tabulate no. of failed studies	Monthly (F)	Operations committee	Principal investigators, program office, study coordinators	Principal investigators	Last day of the month distributed
Quality of studies	Evaluation of studies by center, personnel, and any reading centers	Monthly (M)	Quality control committee	Principal investigators, program office, study coordinators	Principal investigators	Last friday of the month distributed
Technical data distributions	Analysis of technical procedure distributions by center and technicians	Monthly (M)	Quality control committee	Study coordinators	Study coordinators	Last day of the month distributed
Protocol adherence report	Adherence to protocol: No. of technical procedures Equipment maintenance Supervisor checklist	Bimonthly (M)	Operations committee	Principal investigators, quality control committee, study coordinators	Study coordinators	15th of the month following the month distributed
Recruitment by demographics	Analysis of demographics by center	Monthly (F)	Operations committee	Principal investigators, program office, study coordinators	Principal investigators	Calls/meetings of investigators
Data cleanup	By center, identification of outliers and discrepancies in all data sent to coordinating center	Quarterly (M)	Coordinating committee/ quality control committee	Study coordinators	Study coordinators	1 month after date of distribution

* F = sent by fax; M = sent by mail.

often done by means of a pilot study in which the entire protocol is performed on volunteers who are similar demographically to the anticipated cohort. The pilot study should encompass all aspects of the protocol, including all interviews and measured variables, entry and transmission of data to the coordinating center, sending of samples to any reading centers or laboratories involved in the study, processing at the reading centers or laboratories, and creation and distribution of reports by the coordinating center. Some time should then be allowed for implementing refinements to the protocol between the end of the pilot study and the beginning of recruitment.

QUALITY CONTROL

Once data collection begins, the quality control procedures developed as part of the quality assurance process must be implemented. The goal of these quality control procedures is to identify and correct sources of either bias or excessive noise in the data both *during* and *after* data collection. Some of the quality control activities are procedures common to most study designs and will be outlined briefly. The complexity of quality control monitoring introduced by the longitudinal nature of cohort studies will be addressed in more detail. Quality control activities will be divided into those relevant to field centers and those relevant to reading centers or laboratories. The role of the coordinating center will be addressed.

Field center quality control

The quality assurance procedures developed during the planning phase for monitoring technician performance must be implemented. As noted above, it is crucial that reports be timely and that responsibilities for responding to reports and taking actions to correct problems are clear. Frequently, studies establish procedures for identifying problems but fail to establish adequate procedures for ensuring that the problems are resolved. Good quality control reports can be used as an incentive at the field center to encourage top quality work by the data collection staff, serving as a reminder that they are part of a bigger study and that other people outside of their own center are interested in the quality of their work. In cohort studies, typical items to monitor include the following:

Recruitment. Frequent reports should be sent to each center showing how well the site is doing toward meeting recruitment goals, in terms of total numbers recruited and any targets regarding gender or minority enrollment. Data should also be reviewed to ensure that enrolled participants meet all eligibility criteria (e.g., not outside specified age range).

Technician performance. Data should be analyzed regularly to assess technician performance. Evidence that a technician's data deviate significantly from the group as a whole should be reported to the site and investigated by the quality control coordinator. For example, if the overall average systolic blood pressure measured by all technicians during a certain time period was 122.3 mmHg, and the average for one specific technician was 129.7 mmHg, the quality control coordinator should be notified to investigate any potential problems in that technician's technique. In the Cardiovascular Health Study, this type of report uncovered the problem that a certain technician was routinely using a pediatric blood pressure cuff rather than an adult cuff to measure blood pressure on adults, resulting in bias in the data collected by that technician (B. K. Lind, Department of Biostatistics, University of Washington, Seattle, WA, personal communication, 1991). This systematic type of problem can often be corrected using a statistical adjustment. In this example, the effect of using the wrong cuff was quite consistent across all measurements taken with that cuff, so once the size of this effect was determined, a statistical correction factor was applied to the erroneous data.

Monitoring interviewers can be more problematic

Having a quality control supervisor watch an interview is not the best method, as this may impact both the technique that the interviewer uses on that occasion and the participant's responses. Some studies routinely tape all interviews (with the participant's permission), and a random sample is then reviewed by the quality control supervisor for adherence to protocol. In a multicenter study, it is important that tapes are also reviewed by quality control supervisors at other sites to make sure that interviewing techniques are consistent between sites as well as within sites.

In addition to cross-sectional analyses of performance, data should be analyzed by field center and by technician over time to look for drift. Plots and graphs can be especially useful in this type of analysis.

Data clean-up. Field center data should be analyzed routinely for problems such as extreme or inconsistent values. In this effort, longitudinal data are valuable in detecting errors. For example, if a participant's weight is entered as 58 kg one year and 101 kg the next, it is likely that one of these values is an error. Comparing values longitudinally can detect these types of errors, which are easily corrected by asking the field centers to check participant charts.

Replicate measures. Multiple measures taken at one point in time can be used both to identify possible data errors and to calculate more accurate measures of

exposure. For example, multiple blood pressures are often taken during one study visit and averaged in order to give a more accurate value of the participant's blood pressure. Combining multiple measurements by averaging serves to provide more accurate estimates of true data values and to minimize measurement error. For detection of potentially erroneous data, data entry software can be used to flag the data entered if successive values differ by more than a predetermined amount.

For measurements such as blood pressure, it is easy and inexpensive to perform multiple measurements on all participants. In some cohort studies, important outcome or exposure measures are obtained using expensive or invasive technology (e.g., magnetic resonance imaging, blood draws) which cannot reasonably be repeated on all participants. In such situations, replicate measurements can be performed on a subset of participants to quantify variability in the measurement and to obtain a more accurate measure of the variable. For example, in the Sleep Heart Health Study (39), sleep apnea and other sleep-related breathing measurements were obtained using overnight monitoring of participants in their homes. In order to assess how much the values obtained reflected a "first night effect," a second study was performed on a subset of participants. This repeat allowed the Sleep Heart Health Study investigators to determine how accurate their estimates were.

Aspects of field center quality control unique to cohort studies

The longitudinal nature of cohort studies introduces several additional potential quality problems which need to be addressed. These include staff turnover, technician drift, degradation of equipment, technology change, and inconsistencies in participant responses over time.

Problems resulting from staff turnover and technician drift can be addressed in several ways. Established certification requirements will ensure that new staff are trained similarly at all sites, and established requirements for a minimum number of procedures which must be performed each week or month to maintain certification will help minimize drift.

Frequent central training and retraining sessions, held throughout the course of data collection, are powerful tools to use to minimize this problem. These training sessions can have a great impact on data quality in several ways: quality control supervisors from all sites work together and can ascertain that they are all training new staff in comparable ways; all sites can make sure that everyone is interpreting details of the protocol in comparable ways; all staff members get a refresher on protocol details (preferably from someone outside their own site); and staff members can share ideas for retaining participants in the study and for maintaining enthusiasm. In addition to sessions for retraining and recertifying of technicians on study procedures, these sessions should include presentations from study scientists. Scientific sessions on study results are always greatly appreciated by technicians and help to remind them of the importance of their jobs, which can seem mundane on a day-to-day basis.

In order to reduce the costs associated with training, some studies have conducted regional training sessions utilizing computer instruction (16). Other studies have completed certification via telephone for certain procedures (15), or used video cassettes or teleconferencing (30, 32). All of these are valuable tools and are less expensive than central training; however, they all lack some of the benefits of face-to-face training.

Degradation of equipment and technology changes provide another set of unique challenges. It is important that equipment be maintained and calibrated regularly and comparably at all sites, to minimize any measurement error due to equipment. For example, scales should be calibrated regularly and the temperature in freezers used to store samples should be monitored and recorded in a log. In some studies, equipment has a direct impact on data quality. In the Sleep Heart Health Study, overnight sleep data was collected on participants in their homes using a portable monitor. Over time these monitors were subjected to a great deal of wear and tear. Statistics were kept on the proportion of studies scored as "good" or better quality for each monitor, and when a monitor was found to have less than 85 percent "good" or better studies, the sites were notified to have the monitor refurbished. As data collection equipment ages or new technology appears, it can be tempting to replace old equipment with more up-to-date models. However, data must first be collected using both the old and new equipment to ensure that the data are comparable. Changes in either the hardware or software (if applicable) in new equipment can introduce variability or bias in data collected with the new equipment when compared with data collected with the old model.

Data entry and management systems can provide a real challenge in this regard. In cohort studies collecting a broad spectrum of data, a great deal of effort is invested in creating systems for entering, tracking, and managing data. Often the system will become obsolete long before the cohort study ends. The coordinating center or other group responsible for these aspects of the study must decide whether it is worthwhile to invest the effort to move to a more up-to-date and efficient system. An added complication is that in

some cases field staffs are not highly sophisticated computer users, so updating the systems causes complications not only at the coordinating center but for the users at remote sites as well.

A final aspect of field center data quality that must be considered in cohort studies is inconsistent participant response over time. Some data are best collected by self-report, even though the limitations of self-reported data are well-known (23). For example, current smoking status is collected at each visit in many cohort studies, usually by self-report. A certain proportion of participants will give inconsistent answers to smoking questions, e.g., reporting one year that they are current smokers and reporting the next year that they have never smoked. The quality control protocol should recognize that such inconsistencies will occur and have a method in place for handling them. Possible solutions are to recontact participants with inconsistent responses and ask for clarification, to set inconsistent responses to missing values, or, if possible, to use an independent source (such as blood chemistry values) to verify the data.

Quality control of reading centers and laboratories

Many cohort studies send some types of data to central reading centers or laboratories for processing. Reading centers and laboratories present unique quality control challenges because the data processing that they perform adds another source of variability. For example, if echocardiography is performed at field centers, variation is introduced both by the field center technician performing the echocardiogram and by the person at the echocardiogram reading center where the data are interpreted. Often the data processed at such centers or laboratories are the key exposures or outcome variables for the study. Since most cohort studies are interested in measuring real change in exposure variables over time, one goal of quality control procedures should be to assess and minimize extraneous variability in these important exposure measures.

It can be difficult to separate random biologic variability, measurement error, and true change. Therefore, it is essential to get good estimates of biologic variability and measurement error so that true change can be assessed. There are several methods available for doing this.

First, at some point early in the study, a set of samples or tapes should be designated as the "calibration set." This set should be processed at the reading center or laboratory in a blinded fashion at regular intervals, with the results tabulated at the coordinating center. This process serves to identify overall drift over time or the introduction of bias into the data. This calibration set would also be used to train and certify any new readers at the reading center. Second, a carefully designed substudy should be carried out to assess inter- and intrareader variability. In such a study, the coordinating center identifies a set of studies to be reprocessed and assigns multiple readers to each one. These studies are reread by the assigned readers in a blinded fashion, and the results give estimates of the inter- and intrareader variability.

Another important substudy involves having the field center technician perform repeat studies on a subset of participants. Once estimates of inter- and intrareader variability are available, as described above, this type of study allows the estimation of additional variability due to field center technician and biologic variability. (These two aspects of variability usually cannot be estimated separately.) Combining the reader effect and the technician effect gives an estimate of overall measurement error.

Measurement error can lead to either type I or type II errors. For example, when a continuous outcome variable in longitudinal data is measured with error, a type I error can occur. In such a case, a regression analysis looking at the relation between the outcome variable and a set of exposure variables, adjusted for the baseline value of the outcome variable, can show a relation between the observed change in the outcome and the exposure variables even when no association exists between these variables and the true change in the outcome variable (N. David Yanez, Department of Biostatistics, University of Washington, Seattle, WA, personal communication, 1997). High levels of random measurement error can lead to a type II error, obscuring a true relation that may exist.

In a cohort study, quality control analyses of data processed at a reading center or laboratory over time are crucial, even when processes seem automated. One study (the Cardiovascular Health Study) found that cholesterol levels had drifted significantly over a 3-year period even though the laboratory used standard processing and maintained Centers for Disease Control and Prevention certification during that time period. In another case, a study found that certain electrocardiography values jumped during 1 year in the study, due to the hiring of a new reader at the reading center (Mary Ann McBurnie, Department of Biostatistics, University of Washington, Seattle, WA, personal communication, 1996). Often, simple plots and graphs can quickly identify these types of problems.

As in the field centers, staff turnover in a reading center can cause quality problems, in particular problems of drift. When such a problem is identified (using one of the techniques described above, such as the

calibration set or longitudinal plots), several steps are needed to solve the problem and correct the data. First, the reader must undergo a complete retraining/recertification process and must perform up to an acceptable level before processing additional study data. Second, the data processed by the technician should be analyzed to see if a statistical correction is possible. Sometimes drift can be corrected using statistical adjustment. For example, in one acquired immunodeficiency syndrome study, blood CD4 counts drifted over time, so in order to analyze changes in CD4 levels in seropositive individuals, their values had to be adjusted according to the time point and center at which the sample was collected (40).

If statistical adjustment is not possible, complete reprocessing of those data may be required. In the Cardiovascular Health Study, all baseline ultrasound data were reread by the set of readers who processed the follow-up ultrasonograms collected 3 years later, because it was determined that both the software and the reading techniques had changed too much in the intervening time to compare the two sets of data (Lynn R. Shemanski, Department of Biostatistics, University of Washington, Seattle, WA, personal communication, 1998).

Finally, quality control of reading center and laboratory data must include complete tracking procedures. Often, reading centers start a study with little experience processing the volume of data generated by a large, multicenter cohort study. The reading center must be monitored to make sure that it has adequate internal data processing and tracking systems in place, and the coordinating center must make sure that it has adequate systems in place to track data between field centers, reading centers, and the coordinating center. This will minimize data lost due to mishandling, mislabeling, or other problems.

Role of the coordinating center

The coordinating center is generally responsible for performing all of the quality control activities described above. Performing these activities requires a major commitment of personnel, which should be anticipated during planning. The coordinating center must be well-organized, with adequate systems in place to ensure that all reports and quality control analyses are completed and distributed in a timely way (table 2) and that all sites respond to the reports as required. It is also responsible for making sure that study communications are adequate and that all involved investigators are kept up-to-date on any quality control issues that exist. Thus, the value of the coordinating center is not to be underestimated and should receive high priority in its support and maintenance (6), staffed with personnel experienced in data management, quality control, and statistical support.

SUMMARY

As we have presented, it is evident that cohort studies are confronted with their own special, nontrivial issues of quality assurance and quality control. Such studies are typically large-scale designs and involve an extensive amount of data to be collected and processed, the quality of which depends on a variety of factors related to study personnel and equipment. The fact that data are collected over an extended period of time and at several centers greatly increases the magnitude of the data processing task, significantly increasing the likelihood of discrepancies and measurement error in the data.

As presented in tables 1 and 2, the quality assurance and quality control procedures span the entire course of the study and include a multitude of tasks. Such tasks are delegated to various committees and/or are undertaken by participating centers, all of which must take responsibility for understanding, implementing, and following through on all procedures that maximize data quality. The quality of the quality assurance/quality control process is highly correlated with the quality of the communication within and between centers and all researchers. Maintaining standardization of procedures across centers and long-term stability of equipment and analytic procedures are integral components of quality control.

In conclusion, the magnitude of the quality control process in a multicenter longitudinal study should not be underestimated, requiring a significant commitment of study resources. The quality control process is key to the integrity of the study, and an integral part of the design of the study. In a well-designed study, with a good quality control process and dedication to the process by the research team, the validity of the conclusions of the cohort study can be established.

ACKNOWLEDGMENTS

This publication was supported by research grants 5U01-HL53940–04 and NO1HC85079from the National Heart, Lung, and Blood Institute.

REFERENCES

1. Freedland KE, Carney RM. Data management and accountability in behavioral and biomedical research. Am Psychol 1992;47:640–5.
2. Gassman JJ, Owen WW, Kuntz TE, et al. Data quality assur-

ance, monitoring, and reporting. Control Clin Trials 1995;16(2 Suppl):104S-36S.
3. Prud'homme GJ, Canner PL, Cutler JA. Quality assurance and monitoring in the Hypertension Prevention Trial. Hypertension Prevention Trial Research Group. Control Clin Trials 1989;10(3 Suppl):84S-94S.
4. Karrison T. Data editing in a clinical trial. Control Clin Trials 1981;2:15–29.
5. Marinez YN, McMahan CA, Barnwell GM, et al. Ensuring data quality in medical research through an integrated data management system. Stat Med 1984;3:101–11.
6. Bagniewska A, Black D, Molvig K, et al. Data quality in a distributed data processing system: the SHEP Pilot Study. Control Clin Trials 1986;7:27–37.
7. Severe JB, Schooler NR, Lee JH, et al. Ensuring data quality in a multicenter clinical trial: remote site data entry, central coordination and feedback. Psychopharmacol Bull 1989;25: 488–90.
8. Sforza VA. Quality data: what are they? Ann Ist Super Sanita 1994;30:439–43.
9. Hohnloser JH, Puerner F, Soltanian H. Improving coded data entry by an electronic patient record system. Methods Inf Med 1996;35:108–11.
10. Tillotson JL, Gorder DD, DuChene AG, et al. Quality control in the Multiple Risk Factor Intervention Trial Nutrition Modality. Control Clin Trials 1986;7(3 Suppl):66S-90S.
11. Dischinger P, DuChene AG. Quality control aspects of blood pressure measurements in the Multiple Risk Factor Intervention Trial. Control Clin Trials 1986;7(3 Suppl):137S-57S.
12. Cooper GR, Haff AC, Widdowson GM, et al. Quality control in the MRFIT local screening and clinic laboratory. Control Clin Trials 1986;7(3 Suppl):158S-65S.
13. Canner PL, Borhani NO, Oberman A, et al. The Hypertension Prevention Trial: assessment of the quality of blood pressure measurements. Am J Epidemiol 1991;134:379–92.
14. Keltner JL, Johnson CA, Beck RW, et al. Quality control functions of the Visual Field Reading Center (VFRC) for the Optic Neuritis Treatment Trial (ONTT). Control Clin Trials 1993;14:143–59.
15. Anderson MM Jr, Boly LD, Beck RW. Remote clinic/patient monitoring for multicenter trials. Optic Neuritis Study Group. Control Clin Trials 1996;17:407–14.
16. Neaton JD, Duchene AG, Svendsen KH, et al. An examination of the efficiency of some quality assurance methods commonly employed in clinical trials. Stat Med 1990;9:115–24.
17. O'Connor RW, Miller AK. Managing data quality through automation. Toxicology 1987;47:109–18.
18. Fleming TR. Data monitoring committees and capturing relevant information of high quality. Stat Med 1993;12:565–73.
19. Sherbourne CD, Meredith LS. Quality of self-report data: a comparison of older and younger chronically ill patients. J Gerontol 1992;47:S204–11.
20. Habig RL, Thomas P, Lippel K, et al. Central laboratory quality control in the National Cooperative Gallstone Study. Control Clin Trials 1983;4:101–23.
21. Christiansen DH, Hosking JD, Dannenberg AL, et al. Computer-assisted data collection in multicenter epidemiologic research. The Atherosclerosis Risk in Communities Study. Control Clin Trials 1990;11:101–15.
22. Teppo L, Pukkala E, Lehtonen M. Data quality and quality control of a population-based cancer registry. Experience in Finland. Acta Oncol 1994;33:365–9.
23. Correa A, Stewart WF, Yeh HC, et al. Exposure measurement in case-control studies: reported methods and recommendations. Epidemiol Rev 1994;16:18–32.
24. Cook RR. Overview of good epidemiologic practices. J Occup Med 1991;33:1216–20.
25. Lepore P. The 10 essential activities of the quality assurance unit. Qual Assur 1995;4:29–33.
26. Bergman LR. Measurement and data quality in longitudinal research. Eur Child Adolesc Psychiatry 1996;5(Suppl 1): 28–32.
27. Arndt S, Tyrrell G, Woolson RF, et al. Effects of errors in a multicenter medical study: preventing misinterpreted data. J Psychiatr Res 1994;28:447–59.
28. Vantongelen K, Rotmensz N, van der Schueren E. Quality control of validity of data collected in clinical trials. EORTC Study Group on Data Management (SGDM). Eur J Cancer Clin Oncol 1989;25:1241–7.
29. Blumenstein BA. Verifying keyed medical research data. Stat Med 1993;12:1535–42.
30. Johstone FD, Brown MC, Campbell D, et al. Measurement of variables: data quality control. Am J Clin Nutr 1981;34(Suppl 4):804–6.
31. Knatterud GL. Methods of quality control and of continuous audit procedures for controlled clinical trials. Control Clin Trials 1981;1:327–32.
32. van der Putten E, van der Velden JW, Siers A, et al. A pilot study on the quality of data management in a cancer clinical trial. Control Clin Trials 1987;8:96–100.
33. Atherosclerosis Risk in Communities (ARIC) manual 1–12. Chapel Hill, NC: National Heart, Lung, and Blood Institute of the National Institutes of Health, Collaborative Studies Coordinating Center, Department of Biostatistics, University of North Carolina, 1987.
34. Cardiovascular Health Study manual of operations. CHS Coordinating Center, Seattle, Washington, 1989.
35. Sleep Heart Health Study manual of operations. SHHS Coordinating Center, Seattle, Washington, 1995.
36. Mullooly JP. The effects of data entry error: an analysis of partial verification. Comput Biomed Res 1990;23:259–67.
37. Reynolds-Haertle RA, McBride R. Single vs. double data entry in CAST. Control Clin Trials 1992;13:487–94.
38. Cummings J, Masten J. Customized dual data entry for computerized data analysis. Qual Assur 1994;3:300–3.
39. Quan S, Howard BV, Iber C, et al. The Sleep Heart Health Study: design, rationale, and methods. Sleep 1997;20: 1077–85.
40. Muñoz A, Carey V, Saah AJ, et al. Predictors of decline in CD4 lymphocytes in a cohort of homosexual men infected with human immunodeficiency virus. J Acquir Immune Defic Syndr 1988;1:396–404.