Chapter 11
Data Collection and Quality Control

Valid and informative results from clinical trials depend on data that are of high enough quality and sufficiently robust to address the question posed. Such data in clinical trials are collected from several sources—medical records (electronic and paper), interviews, questionnaires, participant examinations, laboratory determinations, or public sources like national death registries. Data elements vary in their importance, but having valid data regarding key descriptors of the population, the intervention, and primary outcome measures is essential to the success of a trial. Equally important, and sometimes a trade-off given limited resources, is having a large enough sample size and number of outcome events to obtain a sufficiently narrow estimate of the intervention effect. Modest amounts of random errors in data will not usually affect the interpretability of the results, as long as there are sufficient numbers of outcome events. However, systematic errors can invalidate a trial's results.

Avoiding problems in the data collection represents a challenge. There are many reasons for poor quality data and avoiding them completely is difficult, so the goal is to limit their amount and, thus, their impact on the trial findings. Many steps can be taken during the planning phase to optimize collection of high quality data. The problems encompass missing data, erroneous (including falsified and fabricated) data, large variability, and long delays in data submission. Even with the best planning, data quality needs to be monitored throughout the trial and corrective actions taken to deal with unacceptable problems. This chapter addresses the problems in data collection, how to minimize collection of poor quality data, and the need for quality monitoring, which includes audits.

Concerted efforts to improve data quality in clinical trials, and to focus on important aspects of quality in large trials, have increased markedly. The International Conference on Harmonisation Good Clinical Practice (ICH-GCP [E6]) guidelines, crafted in the 1990s by a selected group of regulators and industry representatives, defined international ethical and scientific standards for clinical trials [1]. The guidelines cover all the phases of clinical trials from design and conduct to recording and reporting. However, these guidelines are focused on

L.M. Friedman et al., *Fundamentals of Clinical Trials*,
DOI 10.1007/978-3-319-18539-2_11

earlier phase pharmaceutical trials as they are overly complex for many large outcome trials [2, 3]. The roadmap of responsibilities in the ICH-GCP E6 guidance document was most recently revised in 2007 [4], and another revision is in process. Other organizations have issued their own versions of quality assurance guidelines. In 1998, the Society for Clinical Trials issued guidelines for multicenter trials [5]. The oncology community has guidelines issued by the American Society of Clinical Oncology [6] and special standards for pediatric oncology [7]. Others have addressed the specific needs of large trials, including assuring quality without undue regulatory burden. Reports have been published from the 2007 [8] and 2012 [3] Conferences on Sensible Guidelines. A summary of a 2013 meeting of the Clinical Trials Transformation Initiative (CTTI), a public-private partnership founded by the U.S. Food and Drug Administration (FDA) and Duke University, addressed specific issues related to large, simple trials [9]. An article by Acosta et al. [10] discussed the implementation of GCP guidelines in developing countries. The texts by McFadden [11] and Meinert [12] contain detailed descriptions of data collection. Finally, guidance concerning the use of electronic source data in clinical trials has been published by the FDA in 2013 [13] and European Medicines Agency in 2010 [14].

Fundamental Point

During all phases of a study, sufficient effort should be spent to ensure that all data critical to the interpretation of the trial, i.e., those relevant to the main questions posed in the protocol, are of high quality.

The definition of key data depends on trial type and objectives. Baseline characteristics of the enrolled participants, particularly those related to major eligibility measures are clearly key as are primary and important secondary outcome measures, and adverse effects. The effort expended on assuring minimal error for key data is considerable. It is essential that conclusions from the trial be based on accurate and valid data. But fastidious attention to all data is not possible, and in fact can be counterproductive. One approach is to decide in advance the degree of error one is willing to tolerate for each type of data. The key data, as well as certain process information such as informed consent, should be as close to error free as possible. A greater error rate should be tolerated for other data. The confirmation, duplicate testing, and auditing that is done on data of secondary importance should not be as extensive. Perhaps only a sampling of audits is necessary.

In addition to collecting the right data, the method used to collect the data is critical. For some variables, it will be simple collection of numeric information. For other data, the quality depends on carefully constructed questions to assure accurate capture. A well-designed case report form that clearly guides investigators to enter accurate and complete information is critical to the success of the trial.

The data collected should focus on the answers to the questions posed in the protocol. Essential data vary by the type of trial, and they include:

- baseline information, such as inclusion and exclusion criteria that define the population;
- measures of adherence to the study intervention;
- important concomitant interventions;
- primary response variable(s);
- important secondary response variables;
- adverse effects with emphasis on predefined serious events;
- other prespecified response variables.

Data are collected to answer questions about benefits, risks, and ability to adhere to the intervention being tested. Trials must collect data on baseline covariates or risk factors for at least three purposes: (1) to verify eligibility and describe the population studied; (2) to verify that randomization did balance the important known risk factors; and (3) to allow for limited subgroup analyses. Obviously, data must be collected on the primary and secondary response variables specified in the protocol and in some cases tertiary level variables. Some measures of adherence to the interventions specified in the protocol are necessary as well as important concomitant medications used during the trial. That is, to validly test the intervention, the trial must describe how much of the intervention the participant was exposed to and what other interventions were used. Collection of adverse events is challenging for many reasons (see Chap. 12).

Each data element considered should be examined as to its importance in answering the questions. Trialists cannot include every outcome that might be "nice to know." Each data element requires collection, processing, and quality control, as discussed below, and adds to the cost and overall burden of the trial. We think that far too much data are generally collected [15]. Only a small portion is actually used in trial monitoring and publications. Excessive data collection is not only costly but can indirectly affect the quality of the more critical data elements.

Problems in Data Collection

Major Types

There are four major types of data problems discussed here: (1) missing data, (2) incorrect data, (3) excess variability, and (4) delayed submission.

First, incomplete and irretrievably missing data can arise, for example, from the inability of participants to provide necessary information, from inadequate assessment like physical examinations, from laboratory mishaps, from carelessness in completion of data entry, or from inadequate quality control within electronic data management systems. Missing outcome data, for example due to withdrawal of

participant consent or loss to follow-up, can result in unreliable results. When the results of the Anti-Xa Therapy to Lower Cardiovascular Events in Addition to Standard Therapy in Subjects with Acute Coronary Syndrome (ATLAS-ACS 2) trial testing rivaroxaban following acute coronary syndromes were reviewed by an FDA Advisory Committee, drug approval was not recommended in large part due to over 10% of the participants having incomplete follow-up [16]. The percent of missing critical data in a study is considered as one indicator of the quality of the data and, therefore, the quality of the trial.

Second, erroneous data may not be recognized and, therefore, can be even more troublesome than incomplete data. For study purposes, a specified condition may be defined in a particular manner. A clinic staff member may unwittingly use a clinically acceptable definition, but one that is different from the study definition. Specimens may be mislabeled. In one clinical trial, the investigators appropriately suspected mislabeling errors when, in a glucose tolerance test, the fasting levels were higher than the 1-h levels in some participants. Badly calibrated equipment can be a source of error. In addition, the incorrect data may be entered on a form. A blood pressure of 84/142 mmHg, rather than 142/84 mmHg, is easy to identify as wrong. However, while 124/84 mmHg may be incorrect, it is a perfectly reasonable measurement, and the error would not necessarily be recognized. The use of electronic data capture allows automatic checks for data being "out of range" or inconsistent with other data in the participant's record (like diastolic higher than systolic blood pressure). An immediate query can lead to correction right away. The most troublesome types of erroneous data are those that are falsified or entirely fabricated. The pressure to recruit participants may result in alterations of laboratory values, blood pressure measurements, and critical dates in order to qualify otherwise ineligible participants for enrollment [17, 18].

The third problem is variability in the observed characteristics. Variability reduces the opportunity to detect any real changes. The variability between repeated assessments can be unsystematic (or random), systematic, or a combination of both. Variability can be intrinsic to the characteristic being measured, the instrument used for the measurement, or the observer responsible for obtaining the data. People can show substantial day-to-day variations in a variety of physiologic measures. Learning effects associated with many performance tests also contribute to variability. The problem of variability, recognized many decades ago, is not unique to any specific field of investigation [19, 20]. Reports of studies of repeat chemical determinations, determinations of blood pressure, physical examinations, and interpretations of X-rays, electrocardiograms and histological slides indicate the difficulty in obtaining highly reproducible data. People perform tasks differently and may vary in knowledge and experience. These factors can lead to interobserver variability. In addition, inconsistent behavior of the same observer between repeated measurements may also be much greater than expected, though intraobserver inconsistency is generally less than interobserver variability.

Reports from studies of laboratory determinations illustrate that the problem of variability has persisted for almost seven decades. In 1947, Belk and Sunderman [21] reviewed the performance of 59 hospital laboratories on several common

chemical determinations. Using prepared samples, they found that unsatisfactory results outnumbered the satisfactory. Regular evaluation of method performance, often referred to as proficiency testing, is now routinely conducted and required by laboratories in many countries [22, 23]. All laboratories performing measurements for clinical trials should be certified by the Clinical Laboratory Improvement Amendments (CLIA) or a similar agency [24].

Diagnostic procedures that rely on subjective interpretations are not surprisingly more susceptible to variability. One example is radiologists' interpretation of screening mammograms [25]. Nine radiologists read cases with verified cancers, benign, and negative findings in the clinic. Approximately 92% of the mammograms of verified cases were, on average, read as positive. The interradiologist variability was modest. The reading of the negative mammograms showed a substantial interreader variability. In a trial of acute ST segment elevation myocardial infarction, over one-quarter of participants enrolled (and for whom the investigator indicated criteria were met) did not meet inclusion criteria when the electrocardiograms were interpreted by a core laboratory [26]. Since electrocardiogram interpretation in an emergency clinical setting may be less rigorous than in a core laboratory, some degree of disagreement is not surprising.

In another study, the intra- and interreader variability in QT interval measurement on electrocardiograms was estimated by 2 different methods [27]. Eight readers analyzed the same set of 100 electrocardiograms twice 4 weeks apart. Five consecutive complexes were measured. For the more commonly used threshold method, the intrareader standard deviation was 7.5 ms and the interreader standard deviation 11.9 ms. Due to the association between QT prolongation and malignant arrhythmias, the FDA is concerned about drugs that prolong the QT interval by a mean of about 5 ms. Thus, the usual variability in measurement is greater than what is considered a clinically important difference.

Another type of variability is the use of nonstandardized terms. As a result, the ability to exchange, share, analyze, and integrate clinical trial data is limited by this lack of coordination in terms of semantics. Increased attention has been devoted to so-called harmonized semantics [28, 29]. A "universal definition" of myocardial infarction is an attempt to standardize definitions of this event, including harmonizing definitions in clinical trials [30]. In response to the confusion and inconsistency resulting from more than 10 definitions of bleeding used in trials of antithrombotic therapy for coronary intervention, a group of academic leaders and FDA representatives developed a standardized classification of bleeding that has been widely adopted in such trials [31]. Professional societies are becoming engaged in proposing clinical data standards, in large part to establish standard definitions of clinical conditions and outcomes for clinical research [32].

The fourth problem, delayed submission of participant data from the clinical site in multicenter trials, used to be a major issue. However, it has decreased markedly with the onset of electronic data entry (see below).

Minimizing Poor Quality Data

General approaches for minimizing potential problems in data collection are summarized below. Most of these should be considered during the planning phase of the trial. Examples in the cardiovascular field are provided by Luepker et al. [33]. In this section, we discuss design of the protocol and manual, development of data entry tools, training and certification, pretesting, techniques to reduce variability, and data entry.

Design of Protocol and Manual

The same question can often be interpreted in many ways. Clear definitions of entry and diagnostic criteria and methodology are therefore essential. These should be included in the protocol and written so that all investigators and staff can apply them in a consistent manner throughout the trial. Accessibility of these definitions is also important. Even the same investigator may forget how he previously interpreted a question unless he can readily refer to instructions and definitions. A manual of procedures, or the equivalent using an electronic format, should be prepared in every clinical trial. Although it may contain information about study background, design, and organization, the manual is not simply an expanded protocol. In addition to listing eligibility criteria and response variable definitions, it should indicate how the criteria and variables are determined. The manual provides detailed answers to all conceivable "how to" questions, and answers to questions that arise during the trial so they can be documented, shared, and harmonized. Most importantly, the manual should describe the participant visits—their scheduling and content—in detail. Instructions for filling out forms; performing tasks such as laboratory determinations; drug ordering, storing and dispensing; and adherence monitoring must be clear and complete. Finally, recruitment techniques, informed consent, participant safety, emergency unblinding, use of concomitant therapy, and other issues need to be addressed. Updates and clarifications usually occur during the course of a study. These revisions should be made available to every staff person involved in data collection.

Descriptions of laboratory methods or imaging techniques and the ways the results are to be reported also need to be stated in advance. In one study, plasma levels of the drug propranolol were determined by using standardized methods. Only after the study ended was it discovered that two laboratories routinely were measuring free propranolol and two other laboratories were measuring propranolol hydrochloride. A conversion factor allowed investigators to make simple adjustments and arrive at legitimate comparisons. Such adjustments are not always possible.

Development of Forms and Data Entry Tools

Ideally, the study forms, which are increasingly electronic or web-based, should contain all necessary information [12]. If that is not possible, the forms or electronic data entry tools should outline the key information and refer the investigator to the appropriate detailed information. Well-designed tools will minimize errors and variability. Data entry questions and fields should be as clear and well organized as possible, with a logical sequence to the questions. Entry tools should be designed to minimize missing data, for example with inability to proceed until something is entered. To know whether or not a condition is present, one should ask for the answer as "yes" or "no," rather than as a single checkbox if present. This is because the lack of a check mark could mean the condition is not present, it is unknown if it is present, or the question was simply skipped. If it may not be known, then include an "unknown" choice. Questions should be clear, with few "write-in" answers since unstructured text fields will rarely provide helpful information in the typical clinical trial. As little as possible should be left to the imagination of the person completing the form. The questions should elicit the necessary information and little else. Questions that are included because the answers would be "nice to know" are rarely analyzed and may distract attention from the pertinent questions. In several studies where death is the primary response variable, investigators may have an interest in learning about the circumstances surrounding the death. In particular, the occurrence of symptoms before death, the time lapse from the occurrence of such symptoms until death, and the activity and location of the participant at the time of death have been considered important and may help in classifying the cause of death. While this may be true, focusing on these details has led to the creation of extraordinarily complex forms which take considerable time to complete. Moreover, questions arise concerning the accuracy of the information, because much of it is obtained from proxy sources who may not have been with the participants when they died. Unless investigators clearly understand how these data will be used, simpler forms are preferable.

A comprehensive review of the multitude of issues in the design of study forms is presented by Cook and DeMets [34]. They describe the categories of data typically collected in randomized clinical trials: participant identification and treatment assignment; screening and baseline information; follow-up visits, tests, and procedures; adherence to study treatment; adverse experiences; concomitant medication and interventions; clinical outcomes and participant treatment; and follow-up and survival status. Also discussed are mechanisms for data collection and design and review of case report forms.

Training and Certification

There are two types of training for research staff: generic training covering research in general, and training specific to an individual trial. General training includes topics of regulatory requirements, ethics, and basic principles of research and

randomized clinical trials, and this is particularly important for junior investigators and study coordinators (see Chap. 2 for discussion of ethics training). For an individual trial, the training is focused on assuring understanding of the protocol and the ability to faithfully execute it.

It has long been recognized that training sessions for investigators and staff to promote standardization of procedures are crucial to the success of any large study. Whenever more than one person is performing data entry or examining participants, training sessions help to minimize errors. There may be more than one correct way of doing something in clinical practice, but for study purposes, there is only one way. Similarly, the questions on a form should always be asked in the same way. The answer to, "Have you had any stomach pain in the last 3 months?" may be different from the answer to, "You haven't had any stomach pain in the last 3 months, have you?" Even differences in tone or the emphasis placed on various parts of a question can alter or affect the response. Kahn et al. [35] reviewed the favorable impact of training procedures instituted in the Framingham Eye Study. The 2 days of formal training included duplicate examinations, discussions about differences, and the use of a reference set of fundus photographs. Neaton et al. [36] concluded that initial training is useful and should cover the areas of clinic operation, technical measurements, and delivery of intervention. Centralized interim training of new staff is less efficient and can be substituted by regional training, teleconferencing, or web-based approaches.

Mechanisms to verify that clinic staff perform trial procedures and tests the same way, when that may affect trial validity, should be developed. For certain tests, the most reliable interpretation will be using a core laboratory, but even then, standard acquisition of the information at the site must be assured. Mechanisms may include instituting certification procedures for specified types of data collection. If blood pressure is an important outcome in a trial, then there should be standardized procedures for measurement since the approach may have a major impact on the measurement [37]. For certain tests, the people performing these determinations should not only be trained, but also be tested and certified as competent. Periodic retraining and certification are especially useful in long-term studies since people tend to forget, and personnel turnover is common. For situations in which staff must conduct clinical interviews, special training procedures to standardize the approach have been used. In a study of B-mode ultrasonography of the carotid arteries, marked differences in intimal-medial thickness measurements were found between the 13 readers at the reading center [38]. During the 5-year study, the relative biases of readers over time varied, in some cases changing from low to high and vice versa. A sharp increase in average intimal-medial thickness measurements observed toward the end of the study was explained by readers reading relatively high having an increased workload, the hire of a new reader also reading high, and a reader changing from reading low to high.

Pretesting

Pretesting of data entry and procedures is almost always helpful, particularly for variables and formats that have not been used before. Several people similar to the intended participants may participate in simulated interviews and examinations to make sure procedures are properly performed and questions on the forms or screens flow well and provide the desired information. Furthermore, by pretesting, the investigator and staff grow familiar and comfortable with the data entry process. Fictional case histories can be used to check data entry design and the care with which forms are completed. When developing data entry screens, most investigators cannot even begin to imagine the numerous ways questions can be misinterpreted until several people have been given the same information and asked to complete the same form. One explanation for different answers may be due to carelessness on the part of the person completing the data entry. The use of "de-briefing" in the pilot test may bring to light misinterpretations that would not be detected when real participants enter the data. Inadequacies in data entry structure and logic can also be uncovered by use of pretesting. Thus, pretesting reveals areas in which forms might be improved and where additional training might be worthwhile. It also allows one to estimate the time needed to complete data entry, which may be useful for planning, staffing, and budgeting.

De-briefing is an essential part of the training process. This helps people completing data entry understand how the forms are meant to be completed and what interpretations are wanted. Discussion also alerts them to carelessness. When done before the start of the study, this sort of discussion allows the investigator to modify inadequate items on the data entry screens. These case history exercises might be profitably repeated several times during the course of a long-term study to indicate when education and retraining are needed. Ideally, data entry screens should not be changed after the study has started. Inevitably, though, modifications are made, and the earlier the better. Pretesting can help to minimize the need for such modifications.

Techniques to Reduce Variability Including Central Adjudication of Events

Both variability and bias in the assessment of response variables should be minimized through repeat assessment, blinded assessment, or (ideally) both. At the time of the examination of a participant, for example, an investigator may determine blood pressure two or more times and record the average. Performing the measurement without knowing the group assignment helps to minimize bias. A trial of the evaluation of the effect of renal artery denervation on blood pressure illustrates this point. Open-label and less rigorously conducted trials showed a 20 mmHg reduction in systolic blood pressure with the denervation procedure, whereas a larger and

more rigorous sham controlled trial found only a 2 mmHg non-significant effect of the procedure compared with control [39]. In unblinded or single-blinded studies, the examination might be performed by someone other than the investigator, who is blinded to the assignment. In assessing slides, X-rays, images or electrocardiograms, two individuals can make independent, blinded evaluations, and the results can be averaged or adjudicated in cases of disagreement. Independent evaluations are particularly important when the assessment requires an element of judgment and when there is subjectivity in the assessment.

Centralized classification of major health outcomes by blinded reviewers is common in large clinical trials. There are three related objectives: to improve accuracy of event rates and of intervention effect, to reduce bias related to knowledge of intervention assignment (especially in open-label trials), and to improve credibility of results. The common focus of central adjudication is on removing events that are not true events that may create background noise and thus improve the estimate of the true intervention effect. However, it is also possible to identify outcome events that are otherwise missed with centralized clinical events review. For example, myocardial infarction may be difficult to detect around the time of coronary procedures. A systematic central screening for elevated blood cardiac biomarkers substantially increased the number of outcome events detected in the Second Platelet IIb/IIIa Antagonist for the Reduction of Acute Coronary Syndrome Events in a Global Organization Network (PARAGON-B) trial [40].

In an open-label trial, blinding of treatment assignment to adjudicators, as in the case of PROBE (Prospective, Randomized, Open-label, Blinded-Endpoint) design, may reduce bias. It will not eliminate bias, however, since complete blinding is difficult [41] and ascertainment of possible events by the investigator may differ with knowledge of the treatment assignment. It is important to consider aspects other than participant or investigator bias that could impact on event rates in an open-label trial. For example, it is possible that the monthly visits to test for level of anticoagulation (only conducted in the warfarin arm) of the Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) trial [42] could have led to the identification of more events in the warfarin arm.

The central adjudication process may also reduce the variability induced by having a large number of local investigators classifying certain types of events. A critical factor is how well the diagnostic criteria in a trial are specified and communicated to local investigators responsible for the initial classification. Reviews [43, 44] of cardiovascular trials have shown that the event rates and the effects of interventions are only modestly changed when using adjudicated (versus investigator defined) outcomes. And while one might expect the adjudicated results to more clearly demonstrate a treatment benefit of an effective therapy, this was not the case in five of six trials reviewed [43]. It is unclear whether these observations also apply to other disease areas. The FDA encourages the use of standard definitions and of centralized review and classification of critical outcomes [43].

Data Entry

The shift in medical information to an electronic format, both in clinical medicine and in clinical research, has markedly improved data quality and timeliness of data management in clinical trials. Systems are routinely used for data entry as well as for validation of forms and data, document management, and queries and their resolution [45, 46]. Litchfeld et al. [47] compared the efficiency and ease of use of internet data capture with the conventional paper-based data collection system. They reported substantial reductions with the internet-driven approach in terms of time from visit to data entry, time to database release after the last participant visit, and time from a visit to a query being resolved. Seventy-one percent of the sites preferred the web-based approach. Different web-based systems have been developed. Examples include the Validation Studies Information Management System (VSIMS) [48], a system developed for the Childhood Asthma Research (CARE) Network [49], and the Query and Notification System [50].

A variety of approaches are possible to capture and transfer data into the electronic system. The worst case is to first write the data onto paper forms and then transcribe these to the electronic system, since this increases opportunity for transcription error and saves little time. Directly entering the data onto the electronic case report form, or better yet having the data flow directly from the electronic health record, is the goal.

Electronic Source Data

There is a growing opportunity to directly transfer electronic health information into clinical trial databases. Defining when and how this can be done to support pragmatic (and other) clinical trials is a focus of the National Institutes of Health (NIH) Health Systems Collaboratory [51]. Important work is being done to define when and how clinical outcomes can be accurately identified by electronic health records. For example, the FDA's Mini-Sentinel project has developed and validated programs to identify hospitalization for acute myocardial infarction using an algorithm based on the International Classification of Diseases [52]. The FDA has provided guidance for use of electronic source data, emphasizing the same principles that have existed for any source data. This includes that it be "attributable, legible, contemporaneous, original, and accurate (ALCOA)" and also that it meet the regulatory requirements for recordkeeping [13]. Development of integrated electronic systems to direct and manage data flow from various sources is essential for larger trials [45] (Fig. 11.1). These systems can also be leveraged for more efficient direct collection of patient-reported outcomes.

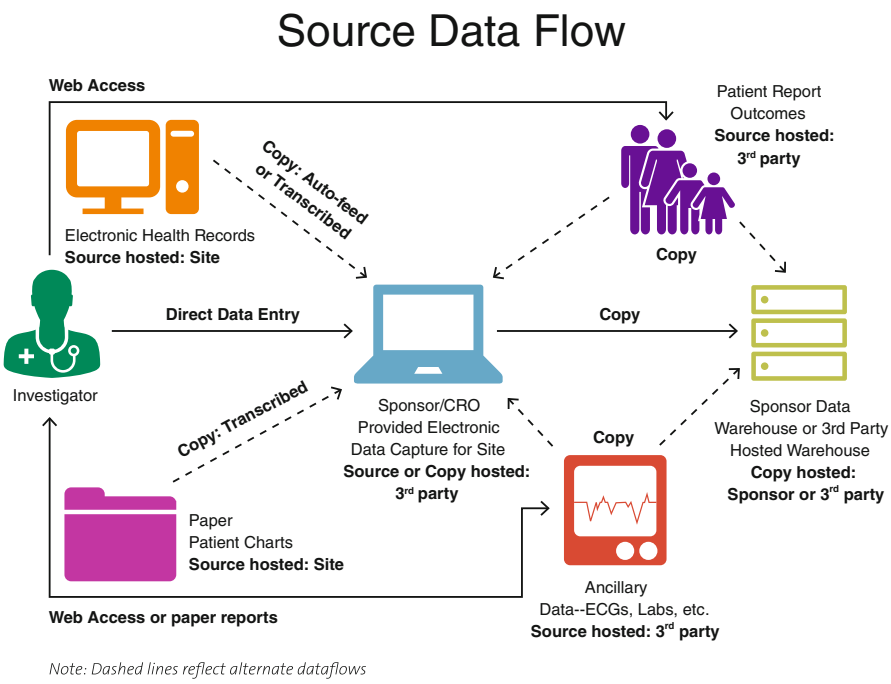

Fig. 11.1 Source dataflow. From Society for Clinical Data Management. eSource Implementation in Clinical Research: A Data Management Perspective: A White Paper [45]

Quality Monitoring

Even though every effort is made to obtain high quality data, a formal monitoring or surveillance system is crucial. When errors are found, this system enables the investigator to take corrective action. Monitoring is most effective when it is current so that when deficiencies are identified, measures can be instituted to fix the problem as early as possible. Additionally, monitoring allows an assessment of data quality when interpreting study results. Numerous procedures, including drug handling and the process of informed consent, can and should be monitored, but monitoring all procedures and study variables will divert resources from more important uses to improve trial quality. Modest amounts of (sometimes unavoidable) random error can be overcome by assuring that there is a robust sample size and number of outcome events. Minimizing missing data, particularly of the primary outcome and major safety outcomes, is crucially important. Monitoring those areas most important to the trial is recommended. This can be done by data quality and consistency checks in the electronic database as well as by on-site review of trial procedures and data.

Monitoring of data quality proves most valuable when there is feedback to the clinic staff and technicians. Once weaknesses and errors have been identified, this

should prompt action to improve the trial through additional training and/or through improving data collection of the problematic variables. Chapter 20 contains several tables illustrating quality control reports. With careful planning, reports can be provided and improvement can be accomplished without unblinding the staff. Investigators need to focus their efforts on those procedures that yield key data, i.e., those on which the conclusions of the study critically depend.

For clinical trials that will form the basis for regulatory decisions, the volume of data is typically very high and the data monitoring very elaborate. Sophisticated clinical trial management systems are used that can integrate electronic data capture, data tracking and management, and other aspects of trial conduct like site performance, pharmacy tracking, adverse event reporting, and payments.

Eisenstein et al. [53, 54] examined ways of reducing the cost of large, phase III trials. The major contributors to the expense are the number of case report form pages, the number of monitoring visits (for comparison of data in source records to the data on trial forms), and the administrative workload. Verification of critical information is important. Limiting the data verification of noncritical data may increase the error rate, but this may have no impact on the overall trial quality as these data are not important to the main findings. There may even be negative impact since limited resources should be focused on where they will make a difference (as outlined in Table 11.1), as opposed to verifying noncritical data. Electronic data entry allows data checks and quality assurance at the time of initial data entry. This can reduce the cost related to traditional "queries" to resolve discrepancies that can be very costly with estimates of more than $100 each. In sensitivity analyses, the authors have shown that the total trial cost could be cut by more than 40% by reducing excessive data collection and verification. Regular site visits to confirm that all case report forms are consistent with patient records is usually excessive. As discussed below, sampling or selective site monitoring would

Table 11.1 Key elements of high quality, randomized clinical outcome trials

Relevant question being addressed
Protocol
– Clear, practical, focused
– Written to avoid "deviations" that are not relevant to "quality"
Adequate number of events to answer question with confidence
Conducted in a general practice setting to make results generalizable
Proper randomization
Reasonable assurance that participants receive (and stay on) assigned intervention
Reasonably complete follow-up and accurate ascertainment of primary outcome (and other key outcomes like death)
Plan for ongoing measurement, feedback, and improvement of quality measures during trial conduct
Safeguards against bias in determining clinically relevant outcomes
Protection of rights of research participants

be more appropriate in most situations. Programs and initiatives like the CTTI [55], as well as FDA guidance on risk-based monitoring [56] and on adverse event reporting [57] are addressing these issues.

Monitoring of Data

During the study, data entered into the system should be centrally checked electronically for completeness, internal consistency, and consistency with other data fields. There should be a system to assure that important source data matches what is in the database, although this can be focused on certain variables and can be supported by selected and focused source-data verification [56]. When the data fields disagree, the group responsible for ensuring consistent and accurate data should assure that a system is in place to correct the discrepancy. Dates and times are particularly prone to error, and systems to minimize missing data are important. Electronic source data systems, especially if they can directly transfer clinical data to an electronic database, can reduce certain types of errors [13].

It may be important to examine consistency of data over time. A participant with a missing leg on one examination was reported to have palpable pedal pulses on a subsequent examination. Cataracts which did not allow for a valid eye examination at one visit were not present at the next visit, without an interval surgery having been performed. The data forms may indicate extreme changes in body weight from one visit to the next. In such cases, changing the data after the fact is likely to be inappropriate because the correct weight may be unknown. The observed differences in measurements may be less dramatic and not obvious. A quality control program based on randomly selected duplicate assessments has been advocated by Lachin [58]. However, the investigator can take corrective action for future visits by more carefully training staff. Sometimes, mistakes can be corrected. In one trial, comparison of successive electrocardiographic readings disclosed gross discrepancies in the coding of abnormalities. The investigator discovered that one of the technicians responsible for coding the electrocardiograms was fabricating his readings. In this instance, correcting the data was possible.

A system should be in place to constantly monitor data completeness and currency to find evidence of missing participant visits or visits that are off schedule, in order to correct any problems. Frequency of missing or late visits may be associated with the intervention. Differences between groups in missed visits may bias the study results. To improve data quality, it may be necessary to observe actual clinic procedures.

Monitoring of Procedures

Extreme laboratory values should be checked. Values incompatible with life such as potassium of 10 mEq/L are obviously incorrect. Other less extreme values (i.e., total cholesterol of 125 mg/dL in male adults in the United States who are not taking

lipid-lowering agents) should be questioned. They may be correct, but it is unlikely. Finally, values should be compared with previous ones from the same participant. Certain levels of variability are expected, but when these levels are exceeded, the value should be flagged as a potential outlier. For example, unless the study involves administering a lipid-lowering therapy, any determination which shows a change in serum cholesterol of 20% or more from one visit to the next should be repeated. Repetition would require saving samples of serum until the analysis has been checked. In addition to checking results, a helpful procedure is to monitor submission of laboratory specimens to ensure that missing data are kept to a minimum.

Investigators doing special procedures (laboratory work, electrocardiogram reading) need to have an internal quality control system. Such a system should include re-analysis of duplicate specimens or materials at different times in a blinded fashion. A system of resubmitting specimens from outside the laboratory or reading center might also be instituted. These specimens need to be indistinguishable from actual study specimens. An external laboratory quality control program established in the planning phase of a trial can detect errors at many stages (specimen collection, preparation, transportation, and reporting of results), not just at the analysis stage. Thus, it provides an overall estimate of quality. Unfortunately, the system most often cannot indicate at which step in the process errors may have occurred.

Recording equipment specific to a trial should be checked periodically. Even though initially calibrated, machines can break down or require adjustment. Scales can be checked by means of standard weights. Factors such as linearity, frequency response, paper speed, and time constant should be checked on electrocardiographic machines. In one long-term trial, the prevalence of specific electrocardiographic abnormalities was monitored. The sudden appearance of a threefold increase in one abnormality, without any obvious medical cause, led the investigator correctly to suspect electrocardiographic machine malfunction.

Monitoring of Drug Handling

In a drug study, the quality of the drug preparations should be monitored throughout the trial. This includes periodically examining containers for possible mislabeling and for proper contents (both quality and quantity). It has been reported that in one trial, "half of the study group received the wrong medication" due to errors at the pharmacy. In another trial, there were concerns about asymmetric mis-allocation of control and experimental drug that turned out to be primarily due to transcription error [59]. Investigators should carefully look for discoloration and breaking or crumbling of capsules or tablets. When the agents are being prepared in several batches, samples from each batch should be examined and analyzed. The actual bottle content of pills should not vary by more than 1% or 2%. The number of pills in a bottle is important to know if pill count will be used to measure participant adherence.

Another aspect to consider is the storage shelf life of the preparations and whether they deteriorate over time. Even if they retain their potency, do changes in odor (as with aspirin) or color occur? If shelf life is long, preparing all agents at one time will minimize variability. Products having a short shelf life require frequent production of small batches. Records should be maintained for study drugs prepared, examined, and used. Ideally, a sample from each batch should be saved. After the study is over, questions about drug identity or purity may arise and samples will be useful.

The dispensing of medication should also be monitored. Checking has two aspects. First, were the proper drugs sent from the pharmacy or pharmaceutical company to the clinic? If the study is double-blind, the clinic staff will be unable to check this. They must assume that the medication has been properly coded. However, in unblinded studies, staff should check to assure that the proper drugs and dosage strengths have been received. In one case, the wrong strength of potassium chloride was sent to the clinic. The clinic personnel failed to notice the error. An alert participant to whom the drug was issued brought the mistake to the attention of the investigator. Had the participant been less alert, serious consequences could have arisen. An investigator has the obligation to be as careful about dispensing drugs as a licensed pharmacist. Close reading of labels is essential, and bar coding can be helpful, as well as documentation of all drugs that are distributed to participants.

Second, when the study is blinded, the clinic personnel need to be absolutely sure that the code number on the container is the proper one. Labels and drugs should be identical except for the code; therefore, extra care is essential. If bottles of coded medication are lined up on a shelf, it is relatively easy to pick up the wrong bottle accidentally. Unless the participant notices the different code, such errors may not be recognized. Even if the participant is observant, he may assume that he was meant to receive a different code number. The clinic staff should be asked to note on a study form the code number of the bottle dispensed and the code number of bottles that are returned by the participants. Theoretically, that should enable investigators to spot errors. In the end, however, investigators must rely on the care and diligence of the staff person dispensing the drugs.

The drug manufacturer assigns lot, or batch, numbers to each batch of drugs prepared. If contamination or problems in preparation are detected, then only those drugs from the problem batch need to be recalled. The use of batch numbers is especially important in clinical trials, since the recall of all drugs can severely delay, or even ruin, the study. When only some drugs are recalled, the study can usually manage to continue. Therefore, the lot number of the drug as well as the name or code number should be listed in the participant's study record.

Audits

There are three general types of audits: routine audits of a random sample of records, structured audits, and audits for cause. Site visits are commonly conducted in long-term multicenter trials. In many non–industry-sponsored trials, a 5–10%

random sample of study forms may be audited for the purpose of verifying accurate transfer of data from hospital source records. This becomes less important if electronic source data can be directly transferred to the database. More complete audits are usually performed in industry-sponsored trials, although there is a movement towards "risk-based monitoring" to focus on critical variables and to customize the intensity of monitoring to the likely benefit of such monitoring [56]. While the traditional model has been for study monitors (or clinical research associates) to visit the sites in order to verify that the entered data are correct, a more appropriate role may be to perform selected source-data verification for critical variables and to spend more time assuring that appropriate systems are in place and training has been performed.

Some investigators have objections to random external data audits, especially in the absence of evidence of scientific misconduct. However, the magnitude of the problems detected when audits occur makes it difficult to take a position against them. Of interest, the FDA has not performed audits of trials sponsored by the National Cancer Institute (NCI), according to a long-standing agreement. It relies on a NCI-sponsored audit program that has been in place since 1982, now known as the Clinical Trials Monitoring Branch of the Cancer Therapy Evaluation Program [60]. A review of 4 cycles of internal audits conducted over an 11-year period by the investigators of the Cancer and Leukemia Group B (CLGB) showed similarities with FDA audits [61]. The deficiency rate (among main institutions) of 28% in the first cycle dropped to 13% in the fourth cycle. Only two cases of major scientific impropriety were uncovered during these on-site peer reviews. Compliance with institutional review board requirements improved over time, as did compliance with having properly signed consent forms. The consent form deficiencies dropped from 18.5% in the first cycle to 4% in the fourth. Although compliance with eligibility improved from 90 to 94%, no changes were noted for disagreement with auditors for treatment responses (5%) and deviations from the treatment protocol (11%). The authors concluded that the audit program had been successful in "pressuring group members to improve adherence to administrative requirements, protocol compliance, and data submission. It has also served to weed out poorly performing institutions."

The NCI National Clinical Trials Network program now has replaced the Cooperative Group Program for overseeing clinical trial activity including quality assurance [60]. Another cooperative group, the National Surgical Adjuvant Breast and Bowel Project, conducted a review of almost 6,000 participant records [62]. The objective was to confirm participant eligibility, disease, and vital status. No additional treatment failures or deaths and only seven cases of ineligible participants were found. The audit was time-consuming and costly and since few discrepancies were found, the authors concluded that routine use of cooperative chart reviews cannot be supported. A similar conclusion was reached in the Global Utilization of Streptokinase and TPA for Occluded Coronary Arteries (GUSTO) trial [63]. Following an audit of all case report forms, the auditors reported only a small percentage of errors and determined that these errors did not change the trial conclusions.

The third type of audit is for cause, i.e., to respond to allegations of possible scientific misconduct. This could be expanded to include any unusual performance pattern such as enrolling participants well in excess of the number contracted for or anticipated. The Office of Research Integrity in the U.S. Department of Health and Human Services promotes integrity in biomedical and behavioral research sponsored by the U.S. Public Health Service at over 7,000 institutions worldwide. It monitors institutional investigations of research misconduct which includes fabrication, falsification or plagiarism in proposing, performing, or reviewing research or in reporting research findings. In a review of 136 investigations resulting in scientific misconduct between 1992 and 2002, only 17 involved clinical trials. The most severe penalty, debarment from U.S. government funding, was applied in six of the cases. Junior employees were often cited and the applied sanction was often a requirement to implement a supervision plan [64, 65].

The FDA conducts periodic audits as well as investigations into allegations of violations of the Federal Food, Drug, and Cosmetic Act through its Office of Criminal Investigations. These may include clinical investigator fraud such as falsifying documentation and enrolling ineligible patients. There were 4,059 FDA inspections in 2013. Most did not justify regulatory action and any corrective action was left to the investigator, with a total of 79 having official action indicated [66].

The quality of any trial is dependent on the quality of its data. Experience has shown that too much data are being collected, much of which are never used for publication or review. As emphasized above, the data collection should be closely linked to the trial objectives and the questions posed in the protocol. The case report form must be carefully constructed to accurately and completely collect the necessary data. Over-collection of data adds to the cost and effort of conducting the trial. Overemphasis on detailed audits of case report forms has similar effects. Moreover, the error rates are often so low that the value of most audits has been questioned, particularly when the errors are “random” in nature. Rather, we should focus our quality control and auditing efforts on key variables. For other variables, samples should be audited with more reliance on statistical quality control procedures. Data collection in clinical trials should be streamlined whenever possible.

References

1. ICH E6. Good clinical practice: Consolidated guideline, Step 5 as of May 1996. Available at http://www.ich.org/products/guidelines/efficacy/efficacy-single/article/good-clinical-practice.html. Accessed January 16, 2015.
2. Yusuf S, Bosch J, Devereaux PJ, et al. Sensible guidelines for the conduct of large randomized trials. *Clin Trials* 2008;5:38-39.
3. Reith C, Landray M, Devereaux PJ, et al. Randomized clinical trials—removing unnecessary obstacles. *N Engl J Med* 2013;369:1061-1065.
4. International Conference on Harmonisation. September 2007. Available at http://www.ich.org. Accessed January 16, 2015.

5. Knatterud GL, Rockhold FW, George SL, et al. Guidelines for quality assurance in multicenter trials: A position paper. *Control Clin Trials* 1998;19:477-493.
6. Zon R, Meropol NJ, Catalano RB, Schilsky RL. American Society of Clinical Oncology statement on minimum standards and exemplary attributes of clinical trial sites. *J Clin Oncol* 2008;26:2562-2567.
7. Devine S, Dagher RN, Weiss KD, Santana VM. Good clinical practice and the conduct of clinical studies in pediatric oncology. *Pediatr Clin North Am* 2008;55:187-208.
8. Baigent C, Harrell FE, Buyse M, et al. Ensuring trial validity by data quality assurance and diversification of monitoring methods. *Clin Trials* 2008;5:49-55.
9. Eapen ZJ, Lauer MS, Temple RJ. The imperative of overcoming barriers to the conduct of large, simple trials. *JAMA* 2014;311:1397-1398.
10. Acosta CJ, Galindo CM, Ochiai RL, et al. Implementation of good clinical practice guidelines in vaccine trials in developing countries. *Vaccine* 2007;25:2852-2857.
11. McFadden E. *Management of Data in Clinical Trials*. Hoboken, New Jersey: John Wiley & Sons, 2007.
12. Meinert CL. *Clinical Trials: Design, Conduct, and Analysis*. Second Edition. New York: Oxford University Press, 2012.
13. FDA Guidance for Industry. Electronic Source Data in Clinical Investigations. Available at http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM328691.pdf. Accessed January 16, 2015.
14. European Medicines Agency. Reflection paper on expectations for electronic source data and data transcribed to electronic data collection tools in clinical trials. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2010/08/WC500095754.pdf. Accessed January 16, 2015.
15. O'Leary E, Seow H, Julian J, Levine M, Pond GR. Data collection in cancer clinical trials: too much of a good thing, *Clin Trials* 2013;10:624-632.
16. Briefing Information for the May 23, 2012 Meeting of the Cardiovascular and Renal Drugs Advisory Committee. Available at http://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/CardiovascularandRenalDrugsAdvisoryCommittee/ucm304754.htm. Accessed January 16, 2015.
17. Neaton JD, Bartsch GE, Broste SK, et al. A case of data alteration in the Multiple Risk Factor Intervention Trial (MRFIT). *Control Clin Trials* 1991;12:731-740.
18. Fisher B, Redmond CK. Fraud in breast-cancer trials. *N Engl J Med* 1994;330:1458-1462.
19. Koran LM. The reliability of clinical methods, data and judgments. Part 1. *N Engl J Med*. 1975;293:642-646.
20. Koran LM. The reliability of clinical methods, data and judgments. Part 2. *N Engl J Med* 1975;293:695-701.
21. Belk WP, Sunderman FW. A survey of the accuracy of chemical analyses in clinical laboratories. *Am J Clin Pathol* 1947;17:853-861.
22. Westgard JO. *Basic Method Validation*. Madison, Wisconsin: Westgard QC, Inc., 2003, pp. 102-103.
23. McPherson RA, Pincus MR (eds.). *Henry's Clinical Diagnosis and Management by Laboratory Methods*. (21st Edition), Philadelphia: Elsevier Saunders, Inc. 2007, pages 4-5.
24. Clinical Laboratory Improvement Amendments (CLIA). Available at http://www.cms.gov/Regulations-and-Guidance/Legislation/CLIA/index.html. Accessed January 16, 2015.
25. Gur D, Bandos AI, Cohen CS, et al. The "laboratory" effect: Comparing radiologists' performance and variability during prospective clinical and laboratory mammography interpretations. *Radiology* 2008;249:47-53.
26. Tjandrawidjaja MC, Fu Y, Al-Khalidi H, et al. Failure of investigator adherence to electrocardiographic entry criteria is frequent and influences clinical outcomes: lessons from APEX-AMI. *Eur Heart J* 2007;28:2850-2857.

27. Panicker GK, Karnad DR, Natekar M, et al. Intra- and interreader variability in QT interval measurement by tangent and threshold methods in central electrocardiogram laboratory. *J Electrocardiol* 2009;42:348-352.
28. Fridsma DB, Evans J, Hastak S, Mead CN. The BRIDG project: A technical report. *J Am Med Inform Assoc* 2008;15:130-137.
29. Weng C, Gennari JH, Fridsma DB. User-centered semantic harmonization: A case study. *J Biomed Inform* 2007;40:353-364.
30. Thygesen K, Alpert JS, Jaffe AS et al. Third universal definition of myocardial infarction. *Eur Heart J* 2012;33:2551-2567.
31. Mehran R, Rao SV, Bhatt DL, et al. Standardized bleeding definitions for cardiovascular clinical trials: a consensus report from the Bleeding Academic Research Consortium. *Circulation* 2011;123:2736-2747.
32. Radford MJ, Heidenreich PA, Bailey SR, et al. American College of Cardiology; American Heart Association Task Force on Clinical Data Standards. ACC/AHA 2007 methodology for the development of clinical data standards: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards. *J Am Coll Cardiol* 2007;49:830-837.
33. Luepker RV, Evans A, McKeigue P, Reddy KS. *Cardiovascular Survey Methods* (Third Edition). Geneva, Switzerland, World Health Organization, 2004.
34. Cook TD, DeMets DL. Data collection and quality control. In Cook TD, DeMets DL (eds.). *Introduction to Statistical Methods for Clinical Trials*. Boca Raton, Florida: Chapman & Hall/CRC, 2007, pp 171-200.
35. Kahn HA, Leibowitz H, Gauley JP, et al. Standardizing diagnostic procedures. *Am J Ophthalmol* 1975;79:768-775.
36. Neaton JD, Duchene AG, Svendson KH, Wentworth D. An examination of the efficacy of some quality assurance methods commonly employed in clinical trials. *Stat Med* 1990;9:115-124.
37. Kandzari DE, Bhatt DL, Sobotka PA, et al. Catheter-based renal denervation for resistant hypertension: rationale and design of the SYMPLICITY HTN-3 Trial. *Clin Cardiol* 2012;35:528-535.
38. Furberg CD, Byington RP, Craven TE. Lessons learned from clinical trials with ultrasound endpoints. *J Intern Med* 1994;236:575-580.
39. Bhatt DL, Kandzari DE, O'Neill WW, et al. A controlled trial of renal denervation for resistant hypertension. *N Engl J Med* 2014;370:1393-1401.
40. Mahaffey KW, Roe MT, Dyke CK, et al. Misreporting of myocardial infarction end points: results of adjudication by a central clinical events committee in the PARAGON-B trial. Second Platelet IIb/IIIa Antagonist for the Reduction of Acute Coronary Syndrome Events in a Global Organization Network Trial. *Am Heart J* 2002;143:242-248.
41. Cohen D. Concerns over data in key dabigatran trial. *BMJ* 2014;349:g4747.
42. Connolly SJ, Ezekowitz MD, Yusuf S, et al. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2009;361:1139-1151.
43. Granger CB, Vogel V, Cummings SR, et al. Do we need to adjudicate major clinical events? *Clin Trials* 2008;5:56-60.
44. Pogue J, Walter SD, Yusuf S. Evaluating the benefit of event adjudication of cardiovascular outcomes in large simple RCTs. *Clin Trials* 2009;6:239-251.
45. Society for Clinical Data Management. eSource Implementation in Clinical Research: A Data Management Perspective. A White Paper. Available at http://www.scdm.org/sitecore/content/be-bruga/scdm/Publications/white%20paper.aspx#sthash.SUHoIyYl.dpufhttp://www.scdm.org/sitecore/content/be-bruga/scdm/Publications/white%20paper.aspx. Accessed January 16, 2015.
46. Reboussin D, Espeland MA. The science of web-based clinical trial management. *Clin Trials* 2005;2:1-2.

47. Litchfield J, Freeman J, Schou H, et al. Is the future for clinical trials internet-based? A cluster randomized clinical trial. *Clin Trials* 2005;2:72-79.
48. Winget M, Kincaid H, Lin P, et al. A web-based system for managing and co-ordinating multiple multisite studies. *Clin Trials* 2005;2:42-49.
49. Schmidt JR, Vignati AJ, Pogash RM, et al. Web-based distributed data management in the Childhood Asthma Research and Education (CARE) Network. *Clin Trials* 2005;2:50-60.
50. Mitchell R, Shah M, Ahmad S, et al. for the Adolescent Medicine Trials Network for HIV/AIDS interventions. A unified web-based query and notification system (QNS) for subject management, adverse events, regulatory, and IRB components of clinical trials. *Clin Trials* 2005;2:61-71.
51. Richesson RL, Hammond WE, Nahm M, et al. Electronic health records based phenotyping in next-generation clinical trials: a perspective from the NIH Health Care Systems Collaboratory. *J Am Med Inform Assoc* 2013;20(e2):e226-31.
52. Cutrona SL, Toh S, Iyer A, et al. Validation of acute myocardial infarction in the Food and Drug Administration's Mini-Sentinel program. *Pharmacoepidemiol Drug Saf* 2012;22:40-54.
53. Eisenstein EL, Collins R, Cracknell BS, et al. Sensible approaches for reducing clinical trial costs. *Clin Trials* 2008;5:75-84.
54. Eisenstein EL, Lemons II PW, Tardiff BE, et al. Reducing the costs of phase III cardiovascular clinical trials. *Am Heart J* 2005;149;482-488.
55. Clinical Trials Transformation Initiative. Available at https://www.ctti-clinicaltrials.org. Accessed January 16, 2015.
56. Guidance for Industry. Oversight of Clinical Investigations—A Risk-Based Approach to Monitoring. Available at http://www.fda.gov/downloads/Drugs/. . ./Guidances/UCM269919.pdf. Accessed January 16, 2015.
57. Guidance for Clinical Investigators, Sponsors, and IRBs. Adverse Event Reporting to IRBs—Improving Human Subject Protection. Available at http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM126572.pdf. Accessed January 16, 2015.
58. Lachin JM. The role of measurement reliability in clinical trials. *Clin Trials* 2004;1:553-566.
59. Alexander JH, Levy E, Lawrence J, et al. Documentation of study medication dispensing in a prospective large randomized clinical trial: experiences from the ARISTOTLE Trial. *Am Heart J* 2013;166:559-65.
60. National Cancer Institute Clinical Trials Monitoring Branch. Available at http://ctep.cancer.gov/branches/ctmb/clinicalTrials/monitoring_coop_ccop_ctsu.htm. Accessed January 16, 2015.
61. Weiss RB, Vogelzang NJ, Peterson BA, et al. A successful system of scientific data audits for clinical trials. *JAMA* 1993;270:459-464.
62. Soran A, Nesbitt L, Mamounas EP, et al. Centralized medical monitoring in phase III clinical trials: the National Surgical Adjuvant Breast and Bowel Project (NSABP) experience. *Clin Trials* 2006;3:478-485.
63. Califf RM, Karnash SL, Woodlief LH. Developing systems for cost-effective auditing of clinical trials. *Control Clin Trials* 1997;18:651-660.
64. Reynolds SM. ORI findings of scientific misconduct in clinical trials and publicly funded research, 1992-2002. *Clin Trials* 2004;1:509-516.
65. Department of Health and Human Services. Public Health Service policies on research misconduct. Final rule. *Federal Register* 2005;70;28370-28400.
66. FDA Inspections Data Dashboard. Available at http://govdashboard.fda.gov/public/dashboards?id=143. Accessed January 16, 2015.