JOURNAL OF CLINICAL RESEARCH BEST PRACTICES

Vol. 2, No. 11, November 2006 "Can You Handle the Truth?"

Are Site Monitoring and Data Cleaning a Waste of Time?

By Norman M. Goldfarb

The objective of a clinical trial is to generate a database of study results. The study sponsor uses this database to support a marketing application (New Drug Application – NDA) with the FDA and other regulatory authorities. For the FDA to consider the study as justification for approval, the study must have adequate statistical power. Statistical power is based on the amount, quality and variability of the data, along with other factors. The FDA reviews the supporting data to confirm that they support the statistical conclusions in the NDA.

Within reasonable bounds, it is not the quality of the data per se that concerns the FDA, but the impact of its quality on statistical power. It is costly to ensure that the quality of the data is adequate, but low-quality data not only may be useless, it may argue against approval of the study drug. Study sponsors thus generate the highest quality data they can. However, the one certain fact about a study database is that it will contain errors; no human process is error-free. Sponsors can reduce the impact of these errors by seeking out competent research sites, conducting thorough site monitoring, and performing rigorous data cleaning – finding possible errors, generating data queries, and correcting the data. Or, they can add subjects to the study to reduce the impact of the errors.

Questions

Automotive, retail and other industries have proven that it is much more effective and less costly to build quality into a product than to inspect the errors out of it.[1] This article thus asks the question:

> Is it more cost-effective to (a) pay competent research sites to generate high-quality data, (b) pay site monitors to inspect the data, (c) pay data managers to generate and process data queries, or (d) pay for some combination of (a), (b) and (c)?

Retail customers care about the quality of products they buy; they want every single unit to be perfect. Statisticians, however, are concerned less about individual pieces of data than in the statistical power of the database as a whole. A larger, lower-quality database is just as useful as a smaller, higher-quality database, provided the size vs. quality tradeoff is made correctly. This article thus asks a second question:

> Is it more cost-effective to obtain statistical power with (a) a smaller, higher-quality database or (b) a larger, lower-quality database?

Allocating resources to the most cost-effective uses saves money that can be used to increase statistical power, accelerate subject recruitment, enhance subject protections, conduct training programs, and invest in new technology.

Regulations, Guidelines and Guidances

Do the regulations require study sponsors to allocate resources between research sites, site monitoring, and data management in particular ways? The answer appears to be "no" – they primarily discuss goals rather than the means used to accomplish those goals. According to the rulebook, study sponsors have substantial flexibility in how they allocate resources to accomplish the regulatory requirements for marketing approval. Over the

years, standard practices have evolved, but it cannot be assumed that the FDA will reject innovations without due consideration.

U.S. federal regulations state:

- The study sponsor is responsible for "ensuring that the investigation(s) is conducted in accordance with the general investigational plan and protocols contained in the IND." (21 CFR § 312.50)
- NDAs are required to include a clinical data section "describing the clinical investigations of the drug, including... a description of the statistical analyses used to evaluate [each controlled clinical] study,... an integrated summary of the data demonstrating substantial evidence of effectiveness for the claimed indications,... [and] case report forms and tabulations." (21 CFR § 314.50)
- "An adequate and well-controlled study [includes] an analysis of the results of the study adequate to assess the effects of the drug. The report of the study should describe the results and the analytic methods used to evaluate them, including any appropriate statistical methods." (21 CFR § 314.126(b))
- "FDA's review of Phases 2 and 3 submissions will... include an assessment of the scientific quality of the clinical investigations and the likelihood that the investigations will yield data capable of meeting statutory standards for marketing approval." (21 CFR § 312.22(a))
- 21 CFR § 312.23 (IND content and format) says nothing about site monitoring or statistical analysis plans.

The FDA "Guideline for the Monitoring of Clinical Investigations" states:

- "A sponsor is responsible for assuring that the data submitted to FDA in support of the safety and effectiveness of a test article are accurate and complete. The most effective way to assure the accuracy of the data submitted to FDA is to review individual subject records and other supporting documents and compare those records with the reports prepared by the investigator for submission to the sponsor. Therefore, during a periodic visit, the monitor should compare a representative number of subject records and other supporting documents with the investigator's reports…" (§ E)
- "The monitor should visit the investigator at the site of the investigation frequently enough to assure that… accurate, complete, and current records are being maintained." (§ D)

ICH guidelines (which are also FDA guidances) state:

- "The sponsor should ensure that the trials are adequately monitored… The determination of the extent and nature of monitoring should be based on considerations such as the objective, purpose, design, complexity, blinding, size, and endpoints of the trial. In general there is a need for on-site monitoring, before, during, and after the trial; however, in exceptional circumstances the sponsor may determine that central monitoring in conjunction with procedures such as investigators' training and meetings, and extensive written guidance can assure appropriate conduct of the trial in accordance with GCP. Statistically controlled sampling may be an acceptable method for selecting the data to be verified. (E6 § 5.18.3)
- "The monitor(s), in accordance with the sponsor's requirements, should ensure that the trial is conducted and documented properly by… verifying that source data/documents and other trial records are accurate, complete, kept up-to-date, and maintained…[and] checking the accuracy and completeness of the CRF entries, source data/documents, and other trial-related records against each other. The monitor specifically should verify that… the data required by the protocol are reported accurately on the CRFs and are consistent with the source data/documents… [and] informing the investigator of any

CRF entry error, omission, or illegibility. The monitor should ensure that appropriate corrections, additions, or deletions are made, dated, explained (if necessary), and initialed by the investigator or by a member of the investigator's trial staff…" (E6 § 5.18.4)

- "The sponsor is responsible for implementing and maintaining quality assurance and quality control systems… to ensure that… data are generated, documents (recorded), and reported in compliance with the protocol, GCP, and the applicable regulatory requirement(s)." (E6 § 5.1.1)
- "The sponsor is responsible for securing agreement from all involved parties to ensure direct access… to all trial-related sites, source data/documents, and reports for the purpose of monitoring and auditing by the sponsor…" (E6 § 5.1.2)
- "Quality control should be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly." (E6 § 5.1.3)
- "The quality assurance and quality control systems implemented to assure the quality of the data should be described in brief [in the clinical study report]. If none were used, this should be stated… Any steps taken at the investigation site or centrally to ensure… the collection of accurate, consistent, complete, and reliable data, such as training sessions, monitoring of investigators by sponsor personnel, instruction manuals, data verification, cross-checking,…, or data audits, should be described. It should be noted whether investigator meetings or other steps were taken to prepare investigators and standardize performance." (E3 § 9.6)
- "The statistical analyses planned in the protocol… should be described [in the clinical study report]…The planned sample size and the basis for it, such as statistical consideration or practical limitations, should be provided…" (E3 § 9.7.2)

The FDA's interest is in "results of the study adequate to assess the effects of the drug." The regulations say nothing about site monitoring, data management, data quality vs. quantity, or statistical analysis plans.

The guidances/guidelines say the monitor should visit the investigative site frequently enough to assure that accurate, complete and current records are being maintained. But, the phrase "frequently enough" includes "never at all" if in-person visits are not required to accomplish the objectives. Statistical sampling (i.e., less than 100%) is allowed. Central (i.e., remote) monitoring is mentioned as a possibility. Site monitoring, data verification, and cross-checking are items in a list of possible quality assurance measures that also includes training and instruction manuals.

According to FDA guidance, sponsors should include in NDA clinical study reports descriptions of the data, statistical methods, and quality measures, but the rules are largely silent about the nature of these elements.

Before implementing a radical interpretation of the rules, sponsors probably want to obtain a blessing from the FDA for their site monitoring, data management, and statistical analysis plans, but, clearly, the regulations and guidances/guidelines are drafted to allow substantial flexibility.

Good and Bad Data

Good data meets the "ALCOA" test – it is attributable, legible, contemporaneous, original and accurate. Bad data consists of data that is unattributable, illegible, recorded after the fact, copied without traceability to the original source, incorrect or missing. Incorrect data is created when study personnel:

- Do not understand the question or how the data is supposed to be entered
- Mis-enter the data (e.g., transpose digits)

- Make a transcription error
- Have bad handwriting
- Use a writing implement that produces illegible or non-permanent text
- Misread an instrument or measurement
- Obtain inaccurate information from a subject or other source
- Invent the data so they don't have to explain why it is missing

Bad data may be identifiable by its implausibility or inconsistency with other data, but it is often indistinguishable from the correct data and slides right through the site monitoring and data cleaning processes.

Data Cost and Quality Model

A spreadsheet model for a hypothetical study can help answer these questions. The "Data Cost and Quality" model at https://www.sitecouncil.org/attachments/0611_Data_Cost.xls includes the following assumptions:

- 500 subjects
- 5 subjects/site
- 200 data points per subject
- All data points are important
- 300% impact of a bad data point on the statistical power of the study; in other words, a bad data point is worse than no data point at all. In the model, a "net good data point" takes into account the impact of bad data points.
- $4,000 site fee per subject
- $20,000 average monitoring cost per site over the course of the study

To measure the impact of site quality, the Data Cost and Quality model includes sub-models for sites of standard, high and low competence, with the assumptions in Table 1.

Table 1: Quality Assumptions

Competence[a]	Standard	High	Low
Site performance cost adjustment[b]	0	-30%	+30%
Site monitoring cost adjustment[c]	0	-25%	+50%
Bad data points[d]	5%	2%	10%
Bad data points corrected by site monitoring[e]	40%	50%	60%
Bad data points detected by data management[f]	40%	20%	70%
Cost per data query[g]	$100	$80	$140
Data points corrected by data management[h]	50%	70%	40%
Mix of sites in study	40%	20%	40%

Notes:
a. Competent sites generate higher-quality data.
b. In addition to generating higher-quality data, competent sites, on average, also enroll more subjects and require less training, monitoring and other sponsor resources.
c. Competent sites make fewer errors, so monitoring is faster, even without considering more the savings from less-than-100% sampling.
d. Missing or incorrect data.

e. It is easier for site monitors to identify errors at less-competent sites because such sites make more obvious errors than more-competent sites. On the other hand, more-competent sites accurately correct more of the errors that monitors find.
f. Competent sites make errors that are harder to detect because they notice and correct the obvious errors themselves.
g. Competent sites are more likely to respond correctly in a timely manner.
h. Competent sites are more likely to accurately correct errors identified in data queries.

Caveats

Assumptions in the Data Cost and Quality model are based on interviews with industry experts, but not hard data; readers may use the model with other assumptions. The model treats all costs as variable. It ignores time spent monitoring regulatory documents, informed consent forms, and drug accountability records, although that time correlates with data quality and the same principles apply. The model considers only the cost of generating data; it does not consider:

- The ramifications of subject population size on other study costs such as study drug and laboratory tests.
- Impact of less competent sites on subject retention and compliance.
- Impact of less competent sites on inspection, regulatory and marketing-approval risk.
- Impact on time required to clean the data. According to an industry survey, the median time from last-subject-visit to database lock is six weeks.[2]
- The human cost of exposing more subjects than absolutely necessary to risks associated with the study. Additional human costs are incurred when site monitoring and data management do not remove as many errors as possible, but even more when low-quality data is generated by low-competency sites.

The conclusions below are based on a hypothetical study and may not apply to all, or any, real-world studies. However, the conclusions certainly raise questions about many common industry practices. The author invites suggestions for refining the model's structure and assumptions.

Conclusions

For the hypothetical study in the Data Cost and Quality model, the following conclusions are clear:

- In a process that includes site monitoring and data cleaning, high-competency sites are the most cost-effective source of high-quality data. Their cost is less than half the cost of high-quality data from low-competency sites. (Total cost/net good data point: High-competency ($29.93), Standard-competency ($44.40), Low-competency ($65.59))
- Site monitoring is not cost-effective, regardless of the competence of the site. (Site cost per net good data point vs. cost of converting bad data points to net good data points with monitoring: High-competency ($14.89 vs. $333.33), Standard-competency ($23.53 vs. $250.00), Low-competency ($37.14 vs. $100.00))
- Site monitoring becomes cost-effective when the average error rate increases from 6% to about 19%. If, instead, the error rate decreases towards zero, the cost of finding an error with site monitoring approaches infinity.
- Site monitoring becomes cost-effective when the cost/subject at a standard-competency site increases to about $22,000.

- The more competent the site, the less cost-effective site monitoring becomes because errors occur less frequently and become more difficult to identify.
- Data cleaning is not cost-effective, regardless of the competence of the site. (Site cost per net good data point vs. cost of converting bad data points to net good data points with data management: High-competency ($14.89 vs. $38.10), Standard-competency ($23.53 vs. $66.67), Low-competency ($37.14 vs. $116.67))
- Data cleaning (after site monitoring) becomes cost-effective when the average error rate increases from 6% to about 26%.
- The less competent the site, the less cost-effective data cleaning becomes because less competent sites require more data cleaning time and correct fewer of the errors found by data managers.
- Data cleaning becomes cost-effective when the cost/subject at a standard-competency site increases to about $15,000.
- The lowest cost net good data points are obtained from high-competency sites without monitoring or data cleaning. ($14.89) In the model, to generate 193,148 net good data points with 100% high-competence sites and no site monitoring or data cleaning costs $1.5 million, a savings of $3.1 million (two-thirds) over the hypothetical study cost of $4.6 million.
- Eliminating all of the low-competency sites reduces the cost of net good data points by 20%, from $49.68 to $39.58, a savings of $4.1 million. If they are replaced with competent sites, the study will also complete faster.
- With a mix of 20% high-competency sites, 40% standard-competency sites, and 40% low-competency sites, it is more cost-effective to add sites and subjects than to monitor all of them or perform data management for all of them. ($26.24 vs. $147.44 vs. $96.83)
- If a bad data point is no worse than not having the data point at all, i.e., a 100% multiplier vs. 300% in the model, the cost of a net good data point goes down slightly for sites ($22.65 vs. $26.24), but increases substantially for site monitoring ($443.31 vs. $147.44) and data management ($290.50 vs. $96.83). The reason is that the multiplier penalizes sites for creating bad data points and rewards site monitors and data managers for fixing them.
- Paying high-competency sites substantially more per subject than low-competency sites is justified by the lower cost of the data they generate. High-competency sites also incur lower human costs and costs for study drug, etc. Higher payments, properly structured, can motivate sites to further improve their data quality, e.g., with training, internal quality assurance, and higher salaries to attract and retain qualified personnel.
- Site monitoring and data management cannot make standard-competency or low-competency sites cost-effective in comparison to high-competency sites. Standard-competency and low-competency sites also incur higher human costs and costs for study drug, etc.
- Site monitors may be more cost-effective if they spend their time helping sites learn how to improve data quality, rather than inspecting the data.
- Data management activities should be restricted to errors that are inexpensive to find and correct, and that have the largest impact on study results.
- Assuming novice sites are among the least competent, sponsors incur a very high cost by working with them. An industry requirement for training and certification of new investigators and study coordinators would reduce this cost.
- Sponsors incur higher-than-necessary costs when they do not know the competency level of their sites. They incur these costs by conducting unnecessary site

monitoring and data management. Establishing long-term relationships with sites minimizes these costs.

- It is more cost-effective to obtain statistical power with a larger, lower-quality database than a smaller, higher-quality database.

Alternatives

Study sponsors can use the Data Cost and Quality model to evaluate different strategies for selecting research sites, allocating resources, and making other decisions that affect data quality.

No sponsor intentionally sets out to find incompetent sites. Many sponsors, however, do not invest adequately in finding, training, motivating, compensating and retaining highly competent sites. Over time, preferred provider relationships can increase the proportion of highly-competent sites. At minimum, sponsors can track – and this seems obvious but apparently is not – remember which sites perform well.

Good sites often complain that sponsors treat them as a commodity, and respond accordingly. Sponsors say they want high-quality data, but their behavior often suggests that what they really want is OK data. Sponsors are more likely to get high-quality data if they reward sites with recognition and monetary compensation for high-quality data. There is nothing unfair or inconsistent about compensating sites for delivering high-quality data. Whereas incentives for subject enrollment may lead to human subject abuses, there is no such risk with incentives for high-quality data.

It is not uncommon for sponsors to "encourage" study managers to accept investigators with more sales potential than research experience. It is not uncommon for study managers to accept "generic" investigators to meet a short-sighted deadline. By using the Data Cost and Quality model, the true cost of these tactics can be revealed.

Technologies such as electronic case report forms (eCRF) and, better yet, electronic source documents (eSource) reduce the rate of errors by minimizing transcription errors and catching errors with real-time edit checks. They implicitly train sites not to repeatedly make the same error. They also reduce the cost of monitoring the data and generate early-warning alerts of problems that can be mitigated by training or closing a site.

100% monitoring of data that has minimal impact on the statistical analysis is probably overkill. Site monitoring time is very expensive: CROs charge roughly $120 per hour for site monitors, including preparation, travel and reporting time, with a net cost per hour onsite of $200 to $400. Adaptive methodologies minimize site monitoring costs and direct available resources to trouble spots. With adaptive sampling, sponsors modify sampling rates based on previous experience with each site within the study and across studies. They can also direct site monitoring resources to data elements that are most prone to error.

Based on its inspections, the FDA is concerned that current site monitoring programs do not, in fact, adequately ensure data quality. The FDA is looking into opportunities for improvement.[3] Study sponsors and the industry as a whole can engage the FDA in discussions about site selection, training, management and monitoring, and data management practices. With the FDA's increasing emphasis on risk management and adaptive (Bayesian) methodologies, it may be more open to innovation than is commonly believed.

Conclusion

We can now answer the two questions asked above:

- It is more cost-effective to pay competent research sites to generate high-quality data than to pay site monitors to inspect the data or pay data managers to generate and process data queries.
- If highly-competent sites are available and properly motivated, it is more cost-effective to obtain statistical power with a smaller, higher-quality database. However, if highly-competent sites are not available, it is more cost-effective to obtain statistical power with a larger, lower-quality database. In either case, reducing the required size of the database with site monitoring and data management is not cost-effective.

These answers do not mean that there should be no site monitoring or data cleaning, only that the functions can be reduced and the resources redeployed to other activities such as finding, training, motivating, compensating and retaining high-quality sites. Data managers can focus less on processing data queries and more on designing and field-testing paper and electronic case report forms that are less prone to error. They can create sophisticated edit checks that identify inconsistencies across multiple subjects and alert data managers to incipient problems. Adaptive methodologies will create new demand for statistical expertise and close study management. More metrics about the studies themselves can be generated and analyzed in sophisticated ways to shorten timelines, improve quality, and reduce cost and risk.

The clinical research industry has evolved in the opposite direction of industries that employ management methods such as Six Sigma and Total Quality Management to produce high-quality, low-cost products and services. A huge amount of time and money is wasted trying to inspect errors out rather than build quality in. No amount of site monitoring and data management can find and correct every error. In fact, FDA investigators find numerous obvious errors that slip through current monitoring programs. If obvious errors are slipping through, how many subtle or invisible errors are there in the rest of the iceberg?

If the objective is high-quality data, the only option is to obtain it from highly competent sites that do not make errors in the first place. Study sponsors could find and develop competent sites and pay them to take responsibility for producing high-quality data. Instead, sponsors pay tens-of-thousands of clinical research associates and data managers to compensate for site deficiencies. Is there any other industry in which the customer is responsible for the quality of the supplier's product?

References

1. "Accelerating the Process of Negotiating Clinical Trial Agreements", Norman M. Goldfarb, Journal of Clinical Research Best Practices, October 2006

2. "Clinical Operations: Accelerating Trials, Allocating Resources & Measuring Performance", Cutting Edge Information, 2006

3. "Good Clinical Practice: A Question & Answer Reference Guide", Section 4: Clinical Monitoring, Mark Mathieu, editor, May 2006

Author

Norman M. Goldfarb is Managing Director of First Clinical Research LLC, a provider of clinical research best practices information, consulting and training services. Contact him at 1.650.465.0119 or ngoldfarb@firstclinical.com.