Central Laboratory Sampling Plans and Quality Control in Clinical Trials

Lemuel A. Moyé, MD, PhD

University of Texas School of Public Health, Health Science Center at Houston, Houston, Texas

ABSTRACT: Central laboratories play an important role in many clinical trials. An important tool in aiding the interpretation of these data has been the use of quality control procedures, primarily achieved by having the central laboratory gather additional data from a subset of the trial participants and therefore obtain a measure of the central laboratory's precision. However, both the additional burden on patients as well as rising clinical trial costs demand that when this trial component is required, it be executed as efficiently as possible. The following analysis explores the relationship between both the number of individuals taking part in the quality assessment and the number of repeated intrapatient measurements on the efficiency of the quality control. This examination reveals that the efficiency of the quality control mechanism for central laboratories is dependent on the sampling scheme used to obtain the specimens, and, in general, a small confidence interval for the variance of reproducibility can be obtained when more than two intrapatient measurements are utilized. Confidence interval lengths are provided for several different combinations of interpatient and intrapatient measurements.

KEY WORDS: *Quality control, central laboratory, variance of reproducibility*

INTRODUCTION

The use of a central laboratory can be important in prospective studies [1]. The analyses these laboratories perform can define entry criteria for patients entering the study, provide an outcome measure for a study, or assess the occurrence of adverse events. Thus, important decisions affecting the trial can be based on the results of the central laboratory. Sometimes corroborative evidence may be required to address the precision and/or the accuracy of the central laboratory in order to incorporate its results into the clinical study with confidence. In several clinical trials [2–4], the central laboratory has undergone a quality control examination. A common tool used in this assessment is repeated measurements of the same sample by the central laboratory. If these measurements require the analysis of blood products, the specimen is divided and forwarded to the central laboratory as different sam-

Address reprint requests to: Lemuel A. Moyé, MD, PhD, University of Texas School of Public Health, 1200 Hermann-Pressler, Houston, TX 77030.

Received October 8, 1990; revised April 2, 1991.

Controlled Clinical Trials 12:761–767 (1991)

655 Avenue of the Americas, New York, New York 10010

0197-2456/91/$3.50

ples. If the measurement requires the reinterpretation of nonperishable material, such as the rereading of echocardiograms for volume determinations, the study can be interpreted once, and then resubmitted for reinterpretation without further burden to the patient. Thus, by requiring the central laboratory to make two intrapatient measurements of identical specimens on n individuals chosen at random from the trial's participants, an estimate of the variance of reproducibility is provided. This quantity is a measure of the central laboratory's precision, and the smaller the variance of reproducibility, the more precise the central laboratory is in performing the assay.

Quality control procedures for central laboratories that yield an estimate of the variance of reproducibility must be performed as efficiently as possible, using the smallest number of analytes or determinations to obtain the best estimate of the variance of reproducibility. The most commonly used sampling frame in this setting is to obtain two measurements on the same sample for each of n individuals selected at random from the trial's participants. This manuscript uses the concept of confidence interval length to determine the optimum combination of interpatient and intrapatient measurements to be used in estimating the variance of reproducibility.

METHODS

Consider a hypothetical prospective clinical trial in which m intrapatient measurements are to be obtained on one specimen obtained from each of n individuals selected at random from the trial's participants. These m intrapatient measurements on n individuals result in nm determinations overall. Let x_{ij} be the jth intrapatient measurement on the ith individual. Using this notation, the quality control data for this variable may be viewed as follows:

Patient 1	$x_{11}\ x_{12}\ x_{13},,,\ x_{1m}$
Patient 2	$x_{21}\ x_{22}\ x_{23},,,\ x_{2m}$
Patient 3	$x_{31}\ x_{32}\ x_{33},,,\ x_{3m}$
	.
	.
	.
Patient n	$x_{n1}\ x_{n2}\ x_{n3},,,\ x_{nm}$

Thus for each individual who has a sample submitted to the central lab as part of the quality control analysis, that sample is divided into m separate specimens to be analyzed by the central laboratory in a masked fashion. For example, if only two specimens per quality control participant are analyzed by the central laboratory for each of the n participants then $m = 2$.

The goal of the following analysis is to determine the smallest values of n and m that ensure the most efficient determination of the central laboratory's precision. One way in which this can be done is to examine the length of the confidence interval obtained for the variance of reproducibility. Since the length of this confidence interval can be constructed as a function of both the number of intrapatient measurements m and number of individuals n, then for a given variance of reproducibility, the values of n and m that yield

the confidence interval of smallest length would be the optimum choices of these quantities.

To begin this analysis, assume x_{ij} has a normal distribution with unknown population mean μ and unknown population variance σ^2. For the ith individual participating in the quality control analysis, we may construct the sample estimate $s_r^2(i)$ of v_r, the variance of reproducibility for the ith individual.

$$s_r^2(i) = \frac{1}{m-1}\sum_{j=1}^{m}(x_{ij} - \bar{x}_i)^2$$

This represents the intrasubject variability for the ith individual only, considering only this individual's collection of m intrapatient measurements. Since the x_{ij} are normal, we also note that

$$(m-1)\frac{s_r^2(i)}{v_r} \sim \chi^2_{m-1}$$

Then, for each of these n individuals involved in the quality control analysis $i = 1, 2, 3, \ldots, n$, one estimates of v_r is available $s_r^2(1)$, $s_r^2(i)$, $s_r^2(3)$, $\ldots$, $s_r^2(n)$. As the information from each individual is independent of that provided by any other individual, we note now that

$$\frac{m-1}{v_r}\sum_{i=1}^{n} s_r^2(i) \sim \chi^2_{n(m-1)} \tag{1}$$

The optimum determination of m and n will be based on the size of the $1 - \alpha$ confidence interval obtainable for v_r. From Eq. 1 it follows that

$$P\left[\chi^2_{n(m-1),\frac{\alpha}{2}} \leq \frac{m-1}{v_r}\sum_{i=1}^{n} s^2_{a/2}(i) \leq \chi^2_{n(m-1),1-\frac{\alpha}{2}}\right] = 1 - \alpha$$

and a confidence interval for v_r, the variance of reproducibility follows.

$$P\left[\frac{(m-1)\sum_{i=1}^{n} s_r^2(i)}{\chi^2_{n(m-1),1-\frac{\alpha}{2}}} \leq v_r \leq \frac{(m-1)\sum_{i=1}^{n} s_r^2(i)}{\chi^2_{n(m-1),\frac{\alpha}{2}}}\right]$$

The focus of our attention will be on the length of the confidence interval $L(m, n, \alpha)$

$$L(m, n, \alpha) = (m-1)\sum_{i=1}^{n} s_r^2(i)\left[\frac{1}{\chi^2_{n(m-1),1-\frac{\alpha}{2}}} - \frac{1}{\chi^2_{n(m-1),\frac{\alpha}{2}}}\right]$$

RESULTS

The determination of $L(m, n, \alpha)$ is computed from a set of χ^2 distribution probability tables. The length of the confidence interval as a function of m and n is particularly useful, allowing the choice of m and n required to obtain a particular precision for the central laboratory's variance of reproducibility. Tables 1 and 2 provide such an analysis for two different measures of s_r^2. For

Table 1 Confidence Interval Widths for the Variance of Reproducibility as a Function of the Number of Subjects and Number of Intrasubject Measurements (sample variance = 10.0)

Number of Subjects	Number of Intrasubject Measurements				
	2	5	10	15	20
50	8.6	4.0	2.6	2.1	1.8
100	5.8	2.8	1.9	1.5	1.3
150	4.7	2.3	1.5	1.2	1.0
200	4.0	2.0	1.3	1.1	0.9
250	3.6	1.8	1.2	0.9	0.8
300	3.3	1.6	1.1	0.9	0.7
350	3.0	1.5	1.0	0.8	0.7
400	2.8	1.4	0.9	0.7	0.6
450	1.8	1.2	0.9	0.8	0.6
500	1.7	1.1	0.9	0.8	0.6

each examination, the type I error α is set at 0.05, and the lengths reported are those of 95% confidence intervals.

Consider the situation in a hypothetical clinical trial in which there are 500 determinations to be used for quality control. For the individuals taking part in quality control, a decision must be made as to the best allocation of these determinations among different individuals and the number of measurements within individuals. If s_r^2 is 10.0 for each participating individual, we may use Table 1 to find that if these 500 determinations are distributed across 250 individuals, each with two measurements, the length of the confidence interval for v_r is 3.6. However, if the 500 determinations are instead allocated over 100 individuals, each with five measurements, the confidence interval length can be reduced from 3.6 to 2.8, a 22% reduction.

An alternative use of the table is to determine the minimum number of determinations required to attain a confidence interval length of 3.0. If only two measurements are attained per individual, 350 individuals are required for a total of 700 determinations. However, if five measurements are allowed per individual, then only 100 individuals are required for a total of 500 de-

Table 2 Confidence Interval Widths for the Variance of Reproducibility as a Function of the Number of Subjects and Number of Intrasubject Measurements (sample variance = 20.0)

Number of Subjects	Number of Intrasubject Measurements				
	2	5	10	15	20
50	17.2	8.0	5.2	4.2	3.6
100	11.6	5.6	3.8	3.0	2.6
150	9.4	4.6	3.0	2.4	2.0
200	8.0	4.0	2.6	2.2	1.8
250	7.2	3.6	2.4	1.8	1.6
300	6.6	3.2	2.2	1.8	1.4
350	6.0	3.0	2.0	1.6	1.4
400	5.6	2.8	1.8	1.4	1.2
450	3.6	2.4	1.8	1.6	1.2
500	3.4	2.2	1.8	1.6	1.2

terminations. Thus, by increasing the number of intrapatient measurements, the total number of determinations and hence the cost of this component of the trial may be reduced.

DISCUSSION

The design of a clinical trial requires that they be executed as efficiently as possible, minimizing both financial cost and burden to the patient. When it has been decided that a central laboratory should be incorporated into the trial design and that quality control procedures should be implemented on that laboratory's measurements, the notion of obtaining the maximum information from its quality measurements must be considered. In selecting patients for the quality control analysis, two extreme options can be considered for collecting measurements from those patients selected to take part in the central laboratory quality control analysis. The first option is to select a number of patients randomly from the quality control cohort of the trial and reinterpret their measure once (one repeat for each of n patients). However, a second point of view might suggest selection of one patient at random from that cohort participating in the quality control analysis and reinterpret that measurement m times. Both of these extreme examples examine an important component of reproducibility of the quality assessment. The purpose of this analysis is to identify the optimum balance of the number of involved patients and the number of intrapatient measurements. The standard against which this balance is found is the length of the confidence interval of the variance of reproducibility.

The previous section demonstrated that different combinations of interpatient and intrapatient measurements lead to confidence intervals of different lengths. The example given previously demonstrates that a 22% reduction in the determinations required can be achieved with no loss of precision in estimating the variance of reproducibility. This is because by increasing the number of intrapatient measurements, an important improvement in the estimation of v_r is obtained, an improvement that is not completely offset by a reduction in the number of individuals participating. It is, of course, the decision of those involved in the design of the study to choose the combination of n and m that leads to a confidence interval of satisfactory length. Figure 1 demonstrates the combination of numbers of interpatient and intrapatient measurements leading to a confidence interval length of 2.5 if the estimated variance of reproducibility is 10.0. There are several assumptions underlying this analysis that must be stated. These assumptions are that the central laboratory is not subject to changes in the assay procedure over time, i.e., there is no random long-term drift. In addition, by assuming that standardization procedures criteria do not change, new reagents are not introduced, and there are no changes in equipment or technique, the variance of reproducibility remains constant over time. These computations also assume the determinations of the central laboratory are not subject to long-term drift over time. If change over time is an issue, then the timing of the repeat measurements is also a factor. Since one of the motivations for a central laboratory is to ensure that measurement precision is uniform for the trial duration, then the intrapatient measurements should be obtained over the duration of the

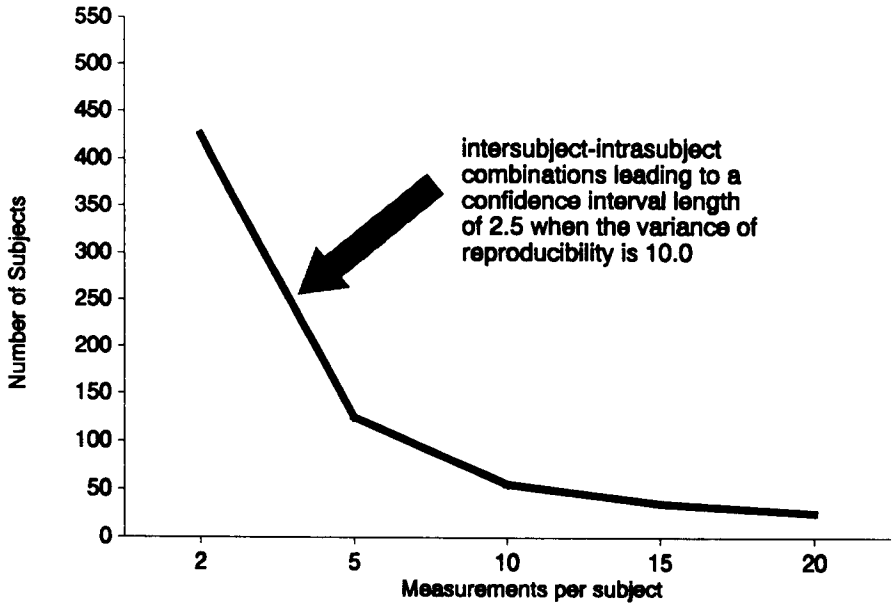

Figure 1 Quality control efficiency: intersubject versus intrasubject measurements.

trial. Statistical tests for examining the variability of the measurement error may be especially useful here.

Many factors must be considered in the allocation of individuals and measurements for quality control including the administration of the process. One of these factors is the availability of adequate samples of replicates. In studies where the measure may be analyzed without being consumed, e.g., the repeat interpretation of ventriculogram or echocardiograms for which the patient supplies no additional analyte, the additional cost is only for the reinterpretation. However, in many other studies, where the specimens to be analyzed are blood products, the amount of serum that is available may severely restrict the number of intrapatient measurements. This is an important consideration in designing quality control mechanisms for central laboratories.

Not all prospective clinical studies require a central laboratory [5], and of those that do, many need not stipulate formal quality control procedures as part of the trial mechanism. However, for those trials that do, the financial costs the trial must support and the burden the patient supplying samples must bear is important. The methods and tables provided here offer some guidance in minimizing the number of patients and intrapatient measurements involved in quality control without losing precision.

The author wishes to thank the Editor-in-Chief and the referees for their many constructive comments, and Ms. Lynda R. Crouch for her help in preparing this manuscript.

REFERENCES

1. Connett JE, Lee WW: Estimation of the coefficient of variation from laboratory analysis in split specimens for quality control in clinical trials. Controlled Clin Trials 11:24–36, 1990
2. MRFIT Research Group: The Multiple Risk Factor Intervention Trial: Quality Control of Technical Procedures and Data Acquisition. Controlled Clin Trials 7:17S–34S, 1986
3. Shulman NB, Ford CE, Hall WD, Blaufox MD, Simon D, Langford HG, Schneider KA: Prognostic value of serum creatinine and the effect of treatment of hypertension on renal function. Results from the HDFP. Hypertension 13(Suppl I):I80–I93
4. The SHEP Cooperative Research Group: Rationale and Design of a Randomized Clinical Trial on Prevention of Stroke in Isolated Systolic Hypertension. J Clin Epidemiol 41:1197–1208, 1988
5. Meinert CL, Tonascia S: Clinical Trials: Design Conduct, and Analysis. New York, Oxford University Press, 1986, pp 36–38