A CONTROL CHART METHOD FOR EVALUATING HEMAGGLUTINATION REAGENT USED IN CHAGAS' DISEASE DIAGNOSIS[1]

Sumie Hoshino-Shimizu,[2] Teresa K. Nagasse-Sugahara,[3] Euclydes A. Castilho,[4] Mario E. Camargo,[5] and Tamio Shimizu[6]

INTRODUCTION

Laboratories that develop or produce their own antigen reagents for serodiagnostic purposes need a practical and reliable quality control method for evaluating the successive batches of standardized reagent, so as to ensure the reproducibility of test results.

Although a large number of statistical models are available for quality control analysis of therapeutic agents or clinical laboratory equipment and procedures *(1, 2)*, very few have been described for the evaluation of serologic reagents. This is probably because such reagents constitute a special category of biological products that measure intricate antibody activities in the sera of infected patients. The measurable output index, known as a titer, results from complex interactions between multiple epitopes (antigenic determinants) of the antigen reagent and a population of polyclonal antibodies whose concentrations will vary depending upon the patient and the stage of the disease. Hence, evaluation of these reagents to ensure that they yield standard, reproducible results demands considerable care.

When we first began preparing hemagglutination reagents for the diagnosis of *Trypanosoma cruzi* infections in our laboratory,[7] we sometimes dealt with anomalous reagents that would give reproducible results with several standard sera, but that proved less sensitive or less specific in routine work than had been indicated by the preliminary evaluations. This problem was better understood when sequential analysis was applied to control the copositivity and conegativity indexes of these reagents by qualitative testing *(3)*. Because

[1] This article will also be published in Spanish in the *Boletín de la Oficina Sanitaria Panamericana*.

[2] Chief, Clinical Immunology Section, Department of Clinical Analysis and Toxicology, Faculty of Pharmaceutical Sciences, University of São Paulo, C.P. 30,786, São Paulo (SP), Brazil.

[3] Biologist, Immunology Laboratory, Institute of Tropical Medicine of São Paulo.

[4] Associate Professor, Department of Preventive Medicine, Faculty of Medicine, University of São Paulo.

[5] Chief, Immunology Laboratory, Institute of Tropical Medicine of São Paulo.

[6] Associate Professor, Department of Industrial Engineering, Polytechnic School, University of São Paulo.

[7] The Immunology Laboratory of The Institute of Tropical Medicine of São Paulo, São Paulo, Brazil.

of the much larger number of serum samples then included in the study of each new batch, it was possible to see that low-titered sera constituted the best indicators for the evaluations.

In most laboratories producing serologic reagents in small batches, quality control is usually performed empirically and based (precariously) upon results obtained with a few standard sera or even with just one pool of sera. This is because sequential analysis, although reliable, requires a relatively large amount of reagent and considerable expense; it therefore seems inappropriate for quality control of reagent batches when the reagent is being produced in limited amounts.

Looking for a better procedure that could provide a practical alternative for evaluating hemagglutination reagents, we investigated the "control chart" or graphic method originally developed by Shewart and cited by others *(4)*. Although employed in industry and in clinical laboratories, it appeared that this method had not previously been applied for evaluating serologic reagent lots. The results of the investigation, which proved promising, are reported here in enough detail so as to permit use of the method studied.

MATERIALS AND METHODS

The hemagglutination test

The reagents were prepared and the hemagglutination tests performed in the manner previously described *(3)*, the reagents being lyophilized and stored at 4°C.

The serum samples

The reagent batches were evaluated with panels of sera made up of serum samples from our laboratory serum bank. These included sera from patients with Chagas' disease, from patients with other unrelated diseases, and from apparently normal individuals. Because it was hard to obtain sera yielding titers below but near the lowest titer considered positive (sera yielding titers of 40 and 80), a special panel was prepared using 23 serum samples obtained from routine work that yielded titers equaling or exceeding 160. The efficiency of these sera in detecting defective reagents was then studied. All of the test sera were preserved in an equal volume of analytical grade glycerin (E. Merck, Darmstadt, Germany) and stored at −20°C. The serum titers obtained were recorded as logarithms of the end-point dilutions, as recommended *(5)*.

The control chart method

This is a simple quality control method that uses a panel of preselected serum samples to evaluate reagent batches in terms of the differences between titers obtained with a test batch and those obtained with a reference reagent. The average standard deviation ($\bar{s}$) of such differences was then plotted on a graph in which a previously defined control limit indicated whether the deviation involved was acceptable or not. The control limit was established on the basis of standard deviations found for reagent batches that had been considered acceptable by a previous statistical study based on sequential analysis, as described *(3)*. It would have been possible, however, to initially establish control limits with batches empirically considered satisfactory.

In practice, the control analysis of a new batch of reagent was accom-

plished by doing serum titrations with panels including 10 reactive and 10 nonreactive serum samples. In order for the reagent to be judged acceptable, the titers obtained with the reactive sera had to yield a standard deviation within the established control limit, and all the nonreactive sera had to yield negative results.

Other statistical methods used

Sequential analysis *(3)* was used to test reagent batches numbered 12 through 18, 25 through 36, and 76 through 79. This method, previously established in our laboratory, determines the acceptability of a reagent according to the extent of false positive or false negative results obtained in qualitative tests of panels including over 150 serum samples, in which about half the sera are positive for Chagas' disease and the remainder are negative. This method involves individual testing of each new reagent batch.

Another method, based on determining the intraclass correlation coefficient (ICC) *(6)*, was also used. This method, which provides an index of agreement derived from analysis of variance, was employed to confirm the relative uniformity of seven reagent batches (those numbered 12 through 18), which were selected for the purpose of deriving a control limit. In addition, the method was also used to confirm the uniformity of reagents 50, 51, and 53, which were subsequently selected for reevaluation of the control limit. The ICC values obtained were procured by testing the subject reagents against a panel of 20 positive and 20 negative sera for *T. cruzi* infection. In the case of reagent batches 12 through 18 the tests were done twice on separate days, and in the case of batches 50, 51, and 53 they were done three times on separate days. ICC values higher than 0.7 were considered acceptable.

Results

A total of 26 batches of Chagas' disease hemagglutination reagent were tested by the control chart method. These included batches that had been rejected, as well as accepted, since 1975 on the basis of sequential analysis.[8]

To determine the reference titer (rT) of each serum sample in the test panels, progressively doubled dilutions of each serum were tested in triplicate against a reference hemagglutination reagent. No differences larger than two dilutions were observed in any of these triple tests. In cases where the three results for a given serum did not coincide, the reference titer was taken to be either the most frequent titer (when two of the three results were the same) or the intermediate titer (when all three results differed).

The seven batches of reagents (numbered 12 through 18) that were selected for the purpose of establishing the limit on the control chart were confirmed as being acceptable by sequential analysis. (Taken together, the seven batches yielded an intraclass correlation coefficient of 0.88.) Each reagent batch was tested against the same panel of 40 serum samples—20 from Chagas' disease patients and 20 from individuals without *T. cruzi* infections. These tests were con-

[8] These 26 batches were numbered as follows: 12–18 (selected to establish the initial control chart limit); 25–36; 50, 51, and 53 (selected to revise the control chart limit); and 76–79. Rejected batches included numbers 28, 30, 33, and 34. All of the other batches were accepted.

ducted in duplicate, the second on a different day from the first, and the results were compared to those obtained with the reference reagent. As may be seen in Figure 1 (part A), about 95% of the titer variations observed were within one dilution of the respective reference titer.

In assessing each reagent batch, the average difference between the observed titers and the reference titer was recorded for each serum, and the standard deviation for the entire panel was calculated for the batch. The arithmetic mean ($\bar{s}$) of the standard deviations of all the batches was then determined *(4)*.

The usually recommended control limit corresponds to three times this mean, or $3\bar{s}$. In our case, the standard deviations observed for the seven batches that had been approved by sequential analysis were 0.467, 0.560, 0.438, 0.494, 0.694, 0.677, and 0.497; and their mean ($\bar{s}$) was 0.547. Three times this latter figure ($3\bar{s}$) was thus 1.64. However, other results obtained by sequential analysis of batches 27, 28, 30, 33, and 34 (the first batch having been accepted and the remaining four rejected) provided another basis for setting the control limit. Since their standard deviations, calculated as shown in Table 1, were 1.34, 1.43, 1.54, 1.60, and 1.42, respectively, it appeared that a lower con-

FIGURE 1. Differences between titers obtained with a reference reagent and those obtained with various reagent batches when tested against 20 sera from Chagas' disease patients. The columns under "A" show the differences found in two tests (on different days) with reagent batches 12 through 18. The columns under "B" show the differences found in three tests (on different days) of improved reagent batches 50, 51, and 53. All of the differences are expressed as $\log_2$ of the endpoint dilution divided by 10 ($\log_2$ 0.1 T − $\log_2$ 0.1 rT).

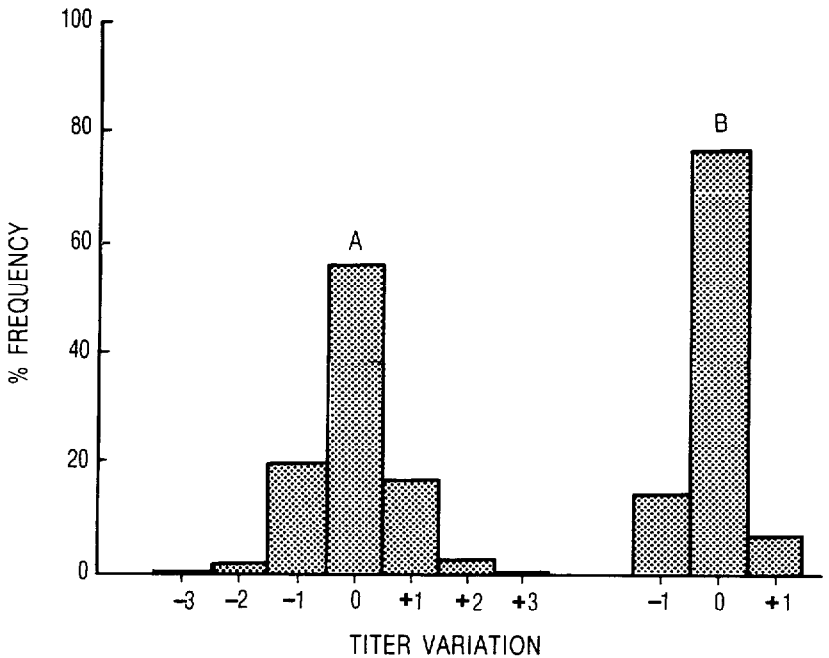


TABLE 1. Assessment of a reagent batch (number 36) by the control chart method. The upper portion of the table shows the results obtained by testing 10 positive sera against the chosen reference hemagglutination reagent and against reagent batch 36. The lower portion shows the equation used to calculate the standard deviation of the reference and test reagent titers, and also shows that batch 36 yielded negative results with all of the 10 negative sera.

Identification numbers of the positive sera tested	Results with reference reagent: Titer (end-point dilution, rT)	Results with reference reagent: Log_2 of: end-point dilution ÷ 10 (log_2 0.1 rT)	Results with reagent batcn 36: Titer (end-point dilution, T)	Results with reagent batcn 36: Log_2 of: end-point dilution ÷ 10 (log_2 0.1 T)	t′ (log_2 0.1 T − log_2 0.1 rT)	$(t')^2$
1	40	2	40	2	0	0
2	80	3	160	4	−1	1
3	80	3	80	3	0	0
4	160	4	160	4	0	0
5	320	5	320	5	0	0
6	160	4	160	4	0	0
7	320	5	320	5	0	0
8	160	4	160	4	0	0
9	1,280	7	2,560	8	+1	1
10	640	6	1,280	7	+1	1

Calculation of $\bar{s}$ for positive sera: $\bar{s} = \sqrt{\sum_{i=1}^{n} (t')^2/n-1}$; $\bar{s} = \sqrt{3/9}$; $\bar{s} = 0.577$ (Accept)

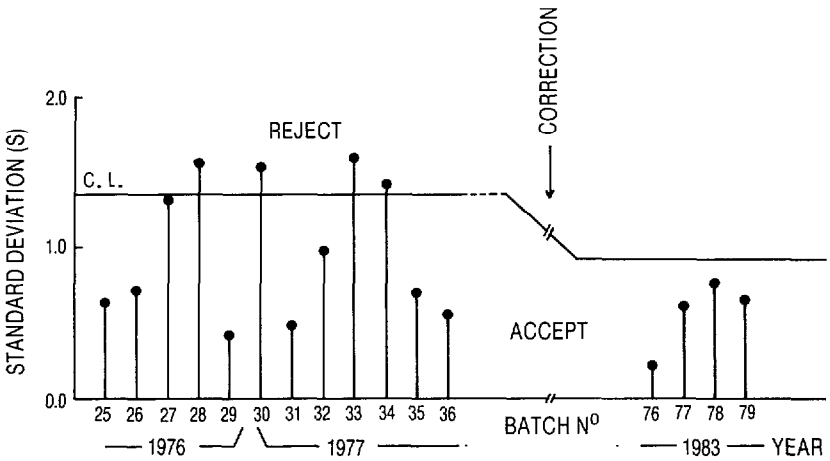
Assessment of results with negative sera: Sera yielding negative results / No. of negative sera tested = 10/10 (Accept)

trol limit between $2\bar{s}$ (1.09) and $3\bar{s}$ (1.64) should be set. Accordingly, $2.5\bar{s}$ was taken as the limit, this corresponding to a variation of 1.37—a limit that would accept batch 27 and reject the others.

Usually, panels of 10 reactive and 10 nonreactive serum samples were employed to test each new reagent batch, as indicated in Table 1. The calculated standard deviation value was then compared to the control limit (Figure 2) to decide whether or not the reagent should be accepted.

Within a few years of the time these procedures had been adopted, however, several improvements were introduced in both preparation and handling of the reagent. As a result, the titer variations were found to be reduced when selected batches (numbered 50, 51, and 53) were tested three times (each time on a different day) against 20 reactive sera, as indicated in Figure 1 (part B). On the basis of this finding, a new control limit was calculated. The standard deviations found for batches 50, 51, and 53 were 0.536, 0.331, and 0.238, respectively, the mean ($\bar{s}$) being 0.368. The new corrected control limit ($2.5\bar{s}$) was thus 0.92 (see the amended limit introduced into Figure 2).

FIGURE 2. A control chart for Chagas' disease hemagglutination reagent batches produced in different years. Batches were accepted or rejected according to the average standard deviation found when the titers they yielded with a panel of 10 negative and 10 positive sera were compared to the titers obtained with a reference reagent. The control limit employed initially (1.37) was later lowered to 0.92 as a result of improved reagent production.



All three of the batches used had been accepted on the basis of prior control chart analysis and had shown a high intraclass correlation coefficient value of 0.96. None of the 10 reagent batches used to set the initial control limit or to reevaluate that limit yielded any false positive or false negative results when tested against panels of 40 sera, 20 from Chagas' disease patients and 20 from uninfected subjects.

A basic step in the foregoing procedures was selection of appropriate serum samples for inclusion in the serum panels to be tested. Before this selection began, it was noted that previous sequential analysis had found low-titered sera (with titers of 40 or 80) to be especially good indicators of poor reagents. However, such low-titered sera are relatively uncommon and hard to get. Therefore, relatively high-titered sera typical of the reactive sera observed in our routine serodiagnosis of Chagas' disease were employed.

Later, to review the appropriateness of this procedure, a panel of 23 serum samples yielding titers of 160 or more was tested against five reagent batches numbered 30 through 34. Three of these batches (30, 33, and 34) had previously been rejected by control chart and sequential analysis evaluations, while two (31 and 32) had been accepted. As already noted, test sera should yield variations (compared to a reference standard) of one dilution or less with a good reagent (see Figure 1, part B), while yielding larger variations with a poor reagent.

To be "useful," the test sera needed to follow this pattern. When tested with the good reagents (31 and 32), all 23 sera yielded titers within one dilution of the reference titer. Conversely, when tested with the poor reagents (30, 33, and 34), 13 of the 23 sera

yielded titers more than one dilution removed from the reference titer for all three reagents (Table 2), and so these 13 sera were considered "useful" for detecting poor reagent batches. Also, six other sera appeared "somewhat useful" because they yielded titers more than one dilution removed from the reference titer with one or two of the poor reagent batches. The four remaining sera were not considered "useful" because they yielded no titers removed from the reference titer by two dilutions or more when tested with the three poor reagent batches. (Sera such as these latter can inadvertently introduce a bias in laboratory work that favors approval of unsatisfactory reagents.)

DISCUSSION AND CONCLUSIONS

This study demonstrates application of the control chart method to evaluate hemagglutination reagents used for Chagas' disease serodiagnosis. Although the method is based on quantitative testing, the statistical techniques involved are simple, and the method is easy to apply. Of course, the choice of reagent testing methods naturally depends on assumptions made about each prevailing situation, but the practical advantages of the control chart technique support its use.

Our study of sera with antibody titers of 160 or more clearly indicates that adequate reagent quality control cannot be ensured by testing the reagent against a few standard sera of this variety, because of the possibility that such sera will belong in the "not useful" category and will be unable to detect defective reagents. In this vein, we found that our few standard sera (prepared by pooling serum samples) had features similar to the sera deemed "not useful," probably because these pooled sera contained high levels of antibodies to most *T. cruzi* epitopes, differing in this respect from most individual sera provided by Chagas' disease patients.

TABLE 2. Results obtained with three poor reagent batches (numbers 30, 33, and 34) and two good reagent batches (numbers 31 and 32) when all five were tested against 23 sera yielding titers of 160 or more with the reference reagent. Thirteen of these sera yielded markedly different titers with the reference reagent than with each of the three poor reagents.

Test sera ($rT \geq 160$) yielding the indicated results		Results with poor reagent batches 30, 33, and 34[a]		Evaluation of test sera
No.	%	Minor titer variations ($t' = 0$ or ± 1)[b]	Major titer variations ($t' > \pm 1$)[b]	
4	17.4	30, 33, 34	—	not useful
2	8.7	33, 34	30	somewhat useful
1	4.3	30, 34	33	somewhat useful
1	4.3	34	30, 33	somewhat useful
2	8.7	30	33, 34	somewhat useful
13	56.6	—	30, 33, 34	useful

[a] Good reagent lots 31 and 32 were also tested. Only narrow titer variations ($t' \pm 1$) were observed with all 23 sera.

[b] $t' = \log_2 0.1\,T - \log_2 0.1\,rT$, where T is the titer obtained with the reagent batch being tested and rT is the titer obtained with the reference reagent.

As brought out previously *(3)*, sequential analysis has shown low-titered sera (with titers of 40 or 80) to be especially good indicators of defective reagent batches. However, the control chart studies reported here show that sera with higher titers can serve as very adequate test samples in serum panels. Fortunately, the preparation of such panels is not difficult, since sera of the latter type can be obtained in the course of normal, routine serodiagnosis of Chagas' disease. Sera that are "not useful" appear to constitute something on the order of 17% of all sera collected in this manner.

In practice, a good panel of sera should have few or no "not useful" sera, consisting entirely of "somewhat useful" and "useful" samples (and perhaps some low-titered sera) in order to ensure sensitivity in detecting poor reagent batches.

In general, the serum samples in a given panel can be replaced by others so long as the others' reference titers and ability to detect unsatisfactory reagents are comparable to those being replaced.

A laboratory wishing to apply the technique described here should start by preparing its own serum panels (testing sera with batches of hemagglutination reagent empirically found good or poor), or else by getting some help with this process from other already-established laboratories. Then, after reference titers are determined for these sera, the control chart limit can be set. Subsequently, simple statistical methods (Student's t test or other suitable procedures) can be used to confirm the validity of the control chart results obtained.

The work reported here was done with panels of serum samples preserved in equal volumes of glycerin at −20°C. This procedure was found to be very appropriate; besides ensuring serum stability, the temperature of −20°C enabled the mixtures of glycerin and serum to remain liquid, so that small aliquots could be removed easily whenever they were needed, without risking the antibody denaturation that tends to result from repeated freezing and thawing.

We recently received a personal communication from Dr. Morris T. Suggs, Director of the Biological Products Program at the United States Center for Infectious Disease in Atlanta, stating that in the United States some 10% to 15% of the serodiagnostic reagents (commercial kits for serodiagnosis) produced each year are unsatisfactory, even those provided by leading manufacturers. He has therefore recommended that one way to obtain good reagents at low cost is to build quality control into production from the beginning. The control chart method described here offers a practical way of assessing successive hemagglutination reagent lots. And, partly because it gives a progressive graphic history of reagent variability, it provides a convenient visual aid for controlling reagent quality.

ACKNOWLEDGMENTS

This investigation was supported in part by Brazil's National Council of Scientific and Technologic Development (*Conselho Nacional de Desenvolvimento Científico e Tecnológico*), Grant No. 222-8-156/80, and by the World Health Organization (WHO/TDR Grant No. T16/81/T8/11(c).)

Summary

Laboratories that produce their own antigen reagents for serodiagnostic purposes need to have practical and reliable ways of testing successive batches of those products so as to ensure the reproducibility of test results. In the case of hemagglutination reagents for the diagnosis of Chagas' disease, the technique of sequential analysis provides adequate quality control. However, sequential analysis requires a relatively large amount of reagent and considerable expense. This article describes another method, a "control chart" technique, that is less elaborate and seems better suited to assessing small reagent batches.

The latter method requires a reference reagent, a panel of some 20 serum samples, and an established limit of variance beyond which the reagent batch under assessment should be rejected. The serum samples should consist half of sera reactive with *T. cruzi* antigen and half of nonreactive sera, and the reactive sera should be "useful" in the sense that they tend to respond differently when tested with a good reagent (such as the reference reagent) than with a poor reagent.

Following this procedure, both the reference reagent and the reagent to be assessed are tested against the serum panel; differences in the titers obtained by the two reagents are noted; the average standard deviation ($\bar{s}$) of these differences is calculated and charted; and if this deviation is less than the previously established control limit, the reagent batch is accepted; otherwise, it is rejected. This method has been used by the Immunology Laboratory at the Institute of Tropical Medicine in São Paulo, Brazil, to test 26 batches of reagent produced at the laboratory since 1975. That experience has shown that sera yielding relatively high titers can be used to detect defective reagents, has indicated that preservation of sera in an equal volume of glycerin and storage at $-20°C$ is effective, and has demonstrated the applicability and usefulness of the control chart method.

References

1 Barnett, R. N. *Clinical Laboratory Statistics (second ed.)*. Little Brown, Boston, 1979.

2 Mack Publishing Company. *Remington's Pharmaceutical Sciences (fourteenth ed.)*. Easton, 1970.

3 Hoshino-Shimizu, S., M. E. Camargo, T. Shimizu, and T. K. Nagasse-Sugahara. A study on the reproducibility of a stable, lyophilized reagent for the Chagas' disease hemagglutination test: Proposals for quality control analysis. *Rev Inst Med Trop São Paulo* 24(2):63–68, 1982.

4 Burr, T. W. *Engineering Statistics and Quality Control*. McGraw-Hill, New York, 1953.

5 White, C. Statistical Methods in Serum Surveys. In: J. R. Paul and C. White (eds.). *Serological Epidemiology*. Academic Press, New York, 1973, p. 19.

6 Winer, B. J. *Statistical Principles in Experimental Design*. McGraw-Hill, New York, 1962.