The Variability of AOAC Methods of Analysis as Used in Analytical Pharmaceutical Chemistry

WILLIAM HORWITZ
Food and Drug Administration, Bureau of Foods, Washington, DC 20204

Pharmaceutical analytical chemistry, which ordinarily deals with the analysis of formulations containing from 0.1 to 100% of active ingredient, uses methods with a reproducibility (between-laboratory variability) of about 2.5% and a repeatability (within-laboratory variability) of about half that amount. The best between-laboratory precision attainable appears to be about 1.0% and within-laboratory precision, about 0.5%. On the basis of the results available, automated methods do not appear to be any more precise than manual methods, although the studies show fewer outlying data points. Replicates (preferably blind ones) should always be conducted in a collaborative interlaboratory study in order to obtain the important information as to whether efforts should be concentrated on improving the method itself or on the performance of laboratories and analysts in applying it.

I like the story told by Lester Chafetz who has a cousin in New York who works at an institute of psychiatry. On his occasional visits there he would meet some of her friends from work. They would ask what he did for a living. He replied that he was an analyst. That made him one of the gang and he could go for an entire evening without it being discovered that he was not their kind of an analyst. Before you discover it for yourselves, I must confess that I have become a food analyst although I spent the first dozen years of my analytical career as a food *and drug* analyst. Because of my affiliation with the Association of Official Analytical Chemists I have maintained a lively interest in the reliability of analytical data, whether it be from foods or drugs. I have been fascinated by the conclusions, summaries, and extrapolations that can be extracted mathematically from published data by popularizers of chemical statistics such as Grant Wernimont, William J. Youden, and, more recently, Gerald Hahn in Chemtech. Their inculcations, sad to say, have been slow to catch on and be put into practice.

There is a great reluctance on the part of analytical chemists to admit that there exists a concept of variability. They take great pride in achieving and reporting the number declared on the label or that the substance under test is 100% pure. In fact, the concept of "by difference" is reported to have been invented by analytical chemists to guarantee achieving perfect analyses. On the other hand, there are also those operators who have never understood the concept of control of variability and whose undisciplined careers have given us the Clinical Laboratories Improvement Act of 1967 and the current effort on the part of the Department of Commerce in behalf of accreditation of laboratories.

Actually we should take as axiomatic that variability exists in all aspects of measurement, whether it be chemical, physical, or biological. An indication of this point is the fact that there are 13,000 members of the American Statistical Association and 26,000 members of the American Society for Quality Control (with some overlap) whose livelihood depends on measuring and interpreting variability. Therefore, for the purpose of this presentation I am going to assume the inevitability of variability and then try to develop a statement of how much variability is inevitable.

For this exercise we are fortunate to have available in the *Journal of the Association of Official Analytical Chemists* every year about a half dozen interlaboratory collaborative studies on drug formulations by many different types of methods. These studies provide us with a bonanza of data which should help us answer the question of how inevitable is our variability. This review has been conducted from the point of view (or bias) of the chemist. From this vantage point, I can dismiss as physically impossible some amazing conclusions of statistical significance ascribed to nonexistent variables by the statistician. In the same breath, I must apologize for the statistically ignorant chemist who so designed his experiment that it permitted such an occurrence. However, I am sure that a statistician, and particularly one who can appreciate the nature of laboratory operations which

produced the data, would be able to extract more subtle and sophisticated conclusions.

In examining these studies, I soon discovered that I had to establish a few ground rules. The AOAC is still grappling with the problem of establishing guidelines for statistical design and interpretation of its collaborative studies. Despite the existence of the helpful manual (1), originally prepared by W. J. Youden of the National Bureau of Standards and later amplified with elementary analysis of variance by E. H. Steiner of Mars, Inc., of England, many of our collaborative studies still do not follow the suggested designs. An analysis of the deficiencies uncovered as a result of the present review of the studies published during the last 5 years will be invaluable in developing such guidelines.

(*1*) *Need for Original Data.* The first of these rules was that I had to eliminate those reports which, in trying to be helpful and concise, buried their variability by reporting the mean of a number of replicates. The AOAC now has a rule that all individual data must be reported and published (or statistically summarized) unless averaging is a specific requirement of the method. (Averaging of a number of replicates to produce a single reported value is a frequent requirement of microbiological assays for antibiotics and vitamins.) Some chemists have the mistaken impression that it is their job to make any method they use look good. This is all too often done by hiding variability through averaging or by discarding apparently aberrant data. These types of manipulations are often disclosed by an abnormally small internal (within-laboratory) variability or are obvious by only a single value reported when replicates are requested. A single value also looks good because it appears to have no variability.

(*2*) *Outliers.* There is a constant argument between chemists and statisticians on the matter of outliers. Some statisticians maintain that an aberrant value (i.e., a value not statistically consistent with the data) should not be discarded because it obviously appeared in the experiment and must be taken into account; therefore, there is a definite probability that such a value may appear again. The chemist, and those chemists like Youden who later became professional statisticians, know how easy it is to make a procedural mistake without realizing it; therefore, they insist that these outliers are the result of involuntary blunders and should not incriminate the "normal" analysts. In this argument, I think that an unbiased observer must say that the chemist is more correct. We have many examples of the analysis of a synthetic sample, where one or two of a dozen analysts are out in left field. If they repeat the work, they almost invariably fall in with the majority. Occasionally they discover and report the cause of the discrepancy. Many of these are the result of errors in weighing and diluting the sample. Some are the result of improperly prepared standard solutions. A copy of the standard curve or a derived value (e.g., absorptivity) should be reported by all participants in a collaborative study as a check for this type of error.

In this connection, we should also note the occasional occurrence of the negative outlier—a laboratory that repeatedly reports less than the normal bias or better than the usual precision. This pattern of data deserves scrutiny from 2 points of view: (*1*) If one laboratory can do this well, why can't the others? (*2*) Are they playing unprofessional games of hiding their variability? The most unfortunate thing about outliers is the fact that their creators are blissfully unaware of their aberrant status. They are more likely to ascribe their values to a random phenomenon than to seek a cause.

In any case, we have accepted the conclusions of those authors who have applied common sense or statistical tests to eliminate outliers. Incidentally, it often makes little difference in the final statistics or in the conclusions regarding the performance of a method whether a single outlier is eliminated in a group of 10 or more. Sometimes, we have substituted a mean for a rejected or missing observation because of the unwieldy mathematics involved in analyzing an unbalanced design if this is not done.

We can, however, put outliers to use. When an excessive number of outliers occur in a study, but with the remaining results showing a satisfactory pattern, usually a correctable cause is

Presented at the 17th Annual Conference on Pharmaceutical Analysis: Pharmaceutical Analysis in an Era of Increasing Regulatory Requirements, University of Wisconsin Extension Services in Pharmacy, Aug. 2, 1977, Lake Delton, WI.

Received July 8, 1977.

present such as directions capable of misinterpretation, improperly prepared standards, or miscalibrated instruments. As a first approximation, we can take a figure of not more than 1 value in 10 that can be discarded by a single laboratory and of not more than 1 laboratory per group of 10 or fraction thereof that may produce outlying data. Such an assumption was made in reviewing aflatoxin methods for referee status in conjunction with a Food and Agriculture Organization (FAO) conference on methods of analysis for contaminants in food (2). The methods of analysis used for parts per billion of aflatoxin in foods, consisting of extraction, adsorptive chromatographic cleanup, and thin layer chromatographic determination, are considerably different from the usual isolation and spectrophotometric determination of parts per hundred of a pharmaceutical.

(3) *Need for Replicates.*—Youden's admonition (1, page 22) to eliminate duplicates has been taken too literally in collaborative studies. Youden actually recommended that the work saved by eliminating duplicates be used to analyze 2 closely similar samples, using the arithmetic differences to remove bias so that only random error (within-laboratory or analyst error) remains. Too many authors read only the first, work-saving portion of the statement. By not following the second portion of the statement, in avoiding the work involved in performing duplicates or by not analyzing 2 closely related samples to be able to plot Youden charts, analysts are eliminating a very valuable source of information regarding a method. When only single determinations per sample are performed in a collaborative study, we cannot estimate the within-laboratory error directly.

In the accompanying tables, the lack of replication is apparent from the omission of an entry under "repeatability." We are defining repeatability here as the standard deviation reflecting within-laboratory or within-analyst error, or the random component of the total error. The term reproducibility is assigned to the standard deviation reflecting all sources of error involving laboratories (which includes analysts, any interactions, and random error) or the total error. In both cases, in order to place all studies on a comparable basis, they are expressed as ratios to the mean values of the active ingredient, which when multiplied by 100 is known as the coefficient of variation. Studies with only a reproducibility value were conducted with only a single analysis per sample per analyst. Since most of the samples within a study were of significantly different composition, they did not satisfy Youden's requirement for closely similar pairs to provide us with the random component of error.

If a laboratory does provide duplicates, there is one more qualification. The duplicates must be independent. For example, aliquots from the same prepared sample solution cannot provide any information with regard to the homogeniety of the sample; aliquots taken at the determinative step cannot provide any information with regard to the variability of all the preceding preparatory steps. Preferably replicates should not be performed as a series but should be randomized so that knowledge of a previous result will not influence the present result. After all, you are looking for variables, not for constants. Practically, however, efficiency demands some semblance of order, if only to avoid the artificial situation of the interaction of high and low samples. Ultimately, we must depend on the professional integrity of analysts to avoid influencing the results of our laboratory work. The best way to provide independent replicates in a collaborative study is to submit duplicates as blind samples to the collaborators. The values reported are now truly independent.

Those studies which provided replicates give us the opportunity of obtaining some very valuable information by comparing the repeatability and the reproducibility terms. The usual case (*see* accompanying tables) is when the repeatability (within-laboratory coefficient of variation) is approximately ½ to ⅔ of the reproducibility (between-laboratory coefficient of variation). Now consider the implications of deviations from this "normal" ratio.

(*a*) *Reproducibility is Very Large.*—A method is unsuitable for the purpose of quality or regulatory control if its variability begins to approach the tolerances permitted in limits and specifications. If your collaborative study shows a reproducibility of ±5% and your specification is 95–105% of the declaration, you would intuitively reject the method as unsuitable. The method must be improved so that the overall analytical variability is only a reasonable fraction of the specification range. This is a case

where analyzing duplicates will be of little value because if the variability is inherent in the method, it will show up just as well among the individual laboratories as in the replicates. (Incidentally, this variability sometimes does not show up in the initiator's laboratory.)

There is one special case where duplicates are invaluable when reproducibility is poor. Laboratories often ascribe their poor performance in a quality control check or in a collaborative study to an inhomogeneous sample. A properly designed study will provide a check on this possibility. Each analyst is requested to weigh out duplicate (or replicate) portions from the submitted specimen, and to analyze at least duplicate aliquots from each of the prepared portions. If the difference between the duplicates (or replicates) on the same sample is small but the means of the duplicates of the 2 samples show a large difference, the sample may be heterogeneous. But if the difference between the duplicates (or replicates) is large and there is little difference between the means of the sample replicates, the discrepancy is more likely to be in the method or in its performance. (A simpler approach, but one which requires considerably more work on the part of the responsible investigator, is to weigh out a known amount of the active ingredient for each sample. The analyst is instructed to use the entire portion for the analysis. There can be no sample variability when the entire sample has been used!)

(*b*) *Repeatability Is Close to Reproducibility.*—In this case all laboratories are performing the method in a consistent manner, and the random component constitutes almost the entire source of variability. If the reproducibility is satisfactory for the intended purpose, there is no need to improve the method.

(*c*) *Repeatability Is Considerably Less than Reproducibility.*—Good repeatability within laboratories but large differences (relatively) between laboratories indicates that the laboratories are conducting the test in a different manner and the source of the discrepancy should be sought. Comparing the absolute values and slopes of standard curves among laboratories may reveal the use of unsuitable or deteriorated reference standards or improperly calibrated instruments. Misinterpretations of the instructions or deliberately changing the method are also frequent sources of variability between laboratories. Prompt follow-up of abnormal values will often uncover and eliminate the source.

In terms of analytical information, omission of independent replicates is a very costly way of saving laboratory time, since the major investment in time and resources has been made in analyzing the first sample; additional samples usually involve the application of only a small fraction of the initial investment. The additional time required for the extra determinations is very cost effective in terms of experimental design.

For completeness, I must point out that Youden states that if the analytical error is less than ⅓ the sampling error, further reduction of the analytical error is of no importance (3). In pharmaceutical work, sampling errors, except possibly in the case of natural products, are usually negligible compared with analytical errors.

Results

The variability of pharmaceutical methods is tabulated by type of method in Tables 1–5. Because of the wide range of units in the original data, all variabilities are expressed as coefficients of variation or relative standard deviation, as it is sometimes called. These values are interpreted as follows:

Repeatability: If an analyst were to replicate the determination a number of times, approximately ⅔ of his results should be within the limits given by the mean, plus or minus the repeatability coefficient of variation (recalculated to the proper units); approximately 95% of his values should be within twice these limits.

Reproducibility: If several laboratories were to analyze a sample a number of times, approximately ⅔ of the laboratory results should be within the limits given by the mean, plus or minus the reproducibility coefficient of variation (recalculated to the proper units); approximately 95% of the values should be within twice these limits.

To present these data compactly, they had to be simplified considerably. For example, Table 1 contains all the methods with a liquid chromatographic step, whether it be partition, ion exchange, or any other type of chromatography where the liquid is the mobile phase, and whether the determinative step is colorimetric, spectrophotometric, or fluorometric. In this table, there is a subset consisting of a single

Table 1. Statistical parameters of methods of analysis for pharmaceutical formulations—liquid chromatographic methods[a]

Ingredient [*JAOAC* ref.]	Magnitude	No. of observations	No. of outliers	Coefficient of variation, %	
				Repeat.	Reprod.
Codeine in APC [(1972) **55**, 142]	1–60 mg	63	1		2.2
Sodium butabarbital tablets [(1972) **55**, 152]	15 mg	52	1	0.66	1.5
Benzthiazide/hydrochlorothiazine [(1972) **55**, 161]	25 mg	120	0	0.9	1.7
Sodium diphenylhydantoin capsules [(1972) **55**, 170]	40%	60	0	0.9	1.6
Nitroglycerin tablets [(1972) **55**, 187]	1% (comp.)	78	6		6.3[b]
	1 (ITA)	78	6		7.8[b]
Estradiol valerate (fluorom.) [(1973) **56**, 86]	4–20 mg/ml	49	0		3.4
Phenylpropanolamine in elixirs (on-col. oxdn) [(1973) **56**, 100]	5–12 mg/5 ml	24 (com.)	0		1.65
		24 (syn.)	0		2.4
Dienestrol tablets (isomerization) [(1973) **56**, 674]	0.1 mg	24	4		3.2
	0.5 mg	24	6		1.8
	pooled	48	10		2.6
Chlorothiazide tablets	70 mg	48	0	0.6	0.8
Methychlorthiazide tablets [(1973) **56**, 677]	1.25 mg	48	0	0.8	2.3
Polythiazide tablets [(1974) **57**, 716]	5–13%	54	0	1.2	2.6
Neostigmine bromide [(1974) **57**, 725]	15 mg/tab.	36	0		2.85
	60 mg/ml	33	3		1.86
Ethinyl estradiol [(1974) **57**, 747]	50–100 μg/tab.	84	6	2.1	2.7
Ephedrine sulfate in sirups [(1975) **58**, 852]	2–5 mg/ml	60	10		2.5
Procaine and propoxycaine solutions	4–19 mg/ml	56	0		1.7
Tetracaine [(1975) **58**, 93]	1.5 mg/ml	14	0		8.8[b]
Bendroflumethiazide	2.4 mg/tab.	48	0	0.4	1.0
Cyclothiazide [(1976) **59**, 90]	2 mg/tab.	48	0	0.91	3.4[b]
Reserpine-rescinnamine alkaloids (fluorom.) [(1976) **59**, 811]	0.15%	40	4		5.6
Mestranol tablets [(1975) **58**, 75]	150 μg	73	7	2.6	2.5
Multicomponent Ion Exchange [(1974) **57**, 741]					
Potassium guaiacol sulfonate:					
mixture 1	20 mg	22	2	2.5	2.8
mixture 2	28 mg	20	2	2.8	4.5
sirup B	9 mg/ml	11	2		1.4
Promethazine:					
mixture 1	25 mg	22	2	2.5	3.2
sirup B	1 mg/ml	10	1		3.9
Codeine:					
mixture 1	20 mg	22	2	2.8	4.9
sirup B	2 mg/ml	10	0		3.6
Pyrilamine:					
mixture 2	20 mg	22	2	1.7	1.8
sirup A	1 mg/ml	11	1		2.8
Dextromethorphan:					
mixture 2	20 mg	18	0	1.7	2.4
mixture 3	25 mg	22	4	3.8	3.6
sirup A	3 mg/ml	11	1		7.0
Phenylpropanolamine:					
mixture 3	25 mg	22	4	2.1	4.6
sirup A	3 mg/ml	11	0 (2)		3.5 (2.8)
Phenylephrine:					
mixture 2	25 mg	22	1	2.5	3.2

(*Continued*)

Table 1 (*Continued*)

Ingredient [*JAOAC* ref.]	Magnitude	No. of observations	No. of outliers	Coefficient of variation, % Repeat.	Reprod.
Chlorpheniramine	25 mg	22	0	3.0	5.8
Glycerol guaiacolate: mixture 3	25 mg	22	2	1.6	1.7
Mean (9 compounds; ion exchange method)				2.4	3.5
Mean (17 studies; ion exchange omitted)				1.1	2.3
Mean (26 compounds)				1.8	2.9

[a] Comp = composite;
ITA = individual tablet analysis;
on-col. oxdn = on-column oxidation;
tab. = tablet;
com. = commercial;
syn. = synthetic;
fluorom. = fluorometric measurement.
[b] Method not adopted.

Table 2. Statistical parameters of methods of analysis for pharmaceutical formulations—gas chromatographic methods

Ingredient [*JAOAC* ref.]	Magnitude	No. of observations	No. of outliers	Coefficient of variation, % Repeat.	Reprod.
Phenol	20 mg/ml	54	0	1.5	3.4
Menthol	15 mg/ml	36	0	1.7	3.0
Methyl salicylate [(1972) **55**, 610]	40 mg/ml	54	0	1.0	3.4
Paraldehyde [(1972) **55**, 166]	20%	55	1	1.0	1.9
Ethanol	4–45%	45	0		2.2
Isopropanol	4–66%	33	0		1.8
Acetone [(1973) **56**, 684]	4%	11	0		2.8
Nikethamide [(1976) **59**, 93]	25%	32	2		2.6
Mean				1.3	2.6
Ratio				0.50	

Table 3. Statistical parameters of methods of analysis for pharmaceutical formulations—extraction with spectrophotometric determination

Ingredient [*JAOAC* ref.]	Magnitude	No. of observations	No. of outliers	Coefficient of variation, % Repeat.	Reprod.
Total trisulfapyrimidines [(1973) **56**, 689]		45	0		1.5
Benztropine mesylate [(1973) **56**, 681]	1 mg/tab.	22	2		2.6
	0.5 mg/tab.	22	0	1.8	2.3
	1 mg/ml	22	2		2.4
Procainamide [(1976) **59**, 807]	80%	96	0	0.8	1.2
	100 mg/ml	24	0	0.7	1.2
Dexamethazone: cream	1 mg/g	30	0 (1)		(visible) 4.8 (3.4)
ointment	0.5 mg/g	30	0 (1)		(UV) 7.1 (5.0)[a]
		30	0		(visible) 2.8
solution [(1974) **57**, 731]	4–25 mg/ml	30	0 (3)		(UV) 7.8 (2.7)[a]
		30	0		(visible) 3.5
		30	0		(UV) 4.7[a]
Nitroglycerin standard [(1972) **55**, 187]	10%	42	0 (6)	1.2 (0.9)	2.2 (1.2)
Mean				1.1	2.5 (2.2)
Ratio				0.44 (0.50)	

[a] Method not adopted.

Table 4. Statistical parameters of methods of analysis for pharmaceutical formulations—automated methods

Ingredient [*JAOAC* ref.]	Amt, mg/tab.	No. of observations	No. of outliers	Coefficient of variation, % Repeat.	Reprod.
Methenamine tablets	325	60	0 (1)	1.5 (1.2)	1.9 (1.7)
std	325	60	0	1.1	1.3
mandelate tablets	500	60	0 (1)	1.3 (1.0)	2.4 (2.3)
+ Na biphosphate [(1973) **56**, 647]	325	60	0	1.1	1.6
Sodium warfarin	2.5–10	180	0	1.0	2.0
Dicumarol [(1973) **56**, 692]	50	60	0	1.2	1.4
Acenocoumarol	4	70	0	1.1	2.1
Phenprocoumon	3	70	1	0.9	2.7
Potassium warfarin [(1975) **58**, 80]	5	70	0	1.3	2.0
Digoxin [(1975) **58**, 70]	0.125	60	0	2.2	3.3
	0.25	60	0	1.3	2.9
	0.5	60	0	1.2	1.8
Reserpine [(1976) **59**, 289]	0.1	70	0	1.6	4.0
	0.25	70	0	1.1	2.9
	0.5	70	1	1.3	2.2
Prednisolone	2.5	60	0	1.2	2.0
	5	60	0	1.1	1.9
Prednisone [(1977) **60**, 27]	2.5	60	0	1.2	1.9
	5	60	0	0.9	1.4
Mean				1.3	2.2
Ratio				0.59	

Table 5. Statistical parameters of methods of analysis for pharmaceutical preparations—miscellaneous methods

Ingredient [*JAOAC* ref.]	Magnitude	No. of observations	No. of outliers	Coefficient of variation, % Repeat.	Reprod.
Polarographic Methods					
Bismuth [(1972) **55**, 155]	10 mg	48	0	1.3	1.8
Glyceryl guaiacolate [(1974) **57**, 756]	50–100 mg	60	0		2.9
Titration					
Nitroglycerin standard [(1972) **55**, 187]	10%	42	0 (3)	1.0 (0.9)	3.5 (1.5)

study (4) with sufficient data for separate analysis involving an ion exchange chromatographic column tested on 9 compounds in various combinations in 5 formulations. In all cases, the original papers should be consulted to obtain the details and complexities of the individual methods.

All of the data were recalculated using a computer program for analysis of variance. In many cases, the statistical parameters may not match those reported in the original paper because most authors simply calculated means and standard deviations without paying attention to a variable number of replicates, or they even included means with individual values. Occasionally means were substituted for missing values to make the data format symmetrical.

In some cases, the data are calculated with and without outliers included. The number of outliers and calculated parameters with that number of outliers omitted are given in parentheses. In arriving at overall estimates for repeatability and reproducibility, I have merely averaged the data, without weighting it or using variances. I know this is statistical heresy, but these estimates have such a large standard deviation of their own that the difference between the results by the 2 mathematical processes is

not worth the effort required to perform the statistical pooling.

The tables do not include a figure for recovery or accuracy. In almost every case, mean recoveries were close to 100%. The studies probably would not have been conducted in the first place if recoveries were not good.

Chromatographic Methods.—Rounded off to one easily remembered significant figure, the within-laboratory coefficient of variation of chromatographic methods is about 1% and the between-laboratory variability (which includes the within-laboratory variability) is about 2%. The estimates for the ion exchange chromatographic method are about double these, 2 and 4%, respectively. This method presumably sacrifices precision in order to obtain the flexibility of operation of a single method for numerous compounds. There is no such thing as a free lunch in analytical chemistry either.

Three methods in this series were rejected by the AOAC. In the case of nitroglycerin, the Associate Referee concluded that he had to control the time between his preparation of the sample and analysis by his collaborators. Ordinarily the method for cyclothiazide would be considered satisfactory, but since the method is almost identical with that for bendroflumethiazide but with a reproducibility coefficient of variation more than 3 times as great, further investigation was requested. In the case of tetracaine, the reproducibility is obviously excessive. The method for reserpine-rescinnamine alkaloids was accepted in view of the nature of the active ingredients.

The ratio for repeatability to reproducibility for all the studies is 0.62, but only 0.48 when the ion exchange chromatographic method is omitted. The ratio for the ion exchange method itself is 0.69. The 0.48 and 0.69 values are the extremes of our self-imposed rule of ratios of approximately ½ to ⅔.

In only a few cases were the number of outliers outside our recommendation of a maximum of 10%. In most of these cases where an excessive number of results were omitted, the reason was reported.

Gas Chromatographic Methods.—There are only 4 studies, involving 8 compounds. The results appear to be in line with those obtained with the liquid chromatographic methods, despite the fact that in most cases the method consists of simply diluting the formulation and injecting into the instrument.

Table 6. Summary of statistical parameters of methods of analysis for pharmaceutical formulations

Type of method	Compounds	Studies	Mean coefficient of variation, % Repeat.	Reprod.
Liquid chromatography	26	18	1.8	2.9
Gas chromatography	8	4	1.3	2.6
Spectrophotometry	5	5	1.1	2.5
Automated	10	7	1.3	2.2
Mean (weighted for compounds)			1.5	2.6
Ratio			0.58	

Spectrophotometric Methods.—The dexamethazone method, which has the highest reproducibility coefficient, is a more complex method than usual, involving an enzyme hydrolysis and extraction prior to the determinative step. A colorimetric determination in this case was superior to ultraviolet spectrophotometry.

Automated Methods.—The data with respect to automated methods are surprising. We have always thought that these methods were certainly more precise, if not more accurate, than manual methods. Even the within-laboratory precision is no better than that of the manual methods! Note, however, that there are far fewer outliers with automated methods than with the other classes.

Miscellaneous Methods.—Table 5 gives the data for 2 polarographic studies and a volumetric method, which are insufficient to draw any conclusions.

Warning.—Although we refer to these statistical parameters of repeatability and reproducibility as characteristic of the method, they are really combination functions of method performance, laboratory performance, and analyst performance. These values are merely average estimates and we cannot say that your laboratory and your analyst will achieve these precisions. They can be used as targets, but the reliability achieved by an analyst in the final analysis depends on skill and experience as well as numerous personal and environmental factors beyond his control. As stated by Shombert (5), "There is *no guarantee* that a naive user of the method will automatically achieve the stated precision. There is an implied *promise* that if he

follows the method carefully in all details he will, in time, with experience, achieve the stated precision."

Tables 1–5 are summarized in Table 6.

Acknowledgment

Kathleen Fominaya mastered the computer program and performed the tedious data input required for the preparation of the tables constituting the basis of this paper.

REFERENCES

(1) Youden, W. J., & Steiner, E. H. (1975) *Statistical Manual of the AOAC*, AOAC, Washington, DC

(2) Schuller, P. L., Horwitz, W., & Stoloff, L. (1976) *JAOAC* 59, 1315–1353

(3) Youden, W. J. (1967) *JAOAC* 50, 1007–1013

(4) Smith, D. J. (1974) *JAOAC* 57, 741–746

(5) Shombert, G., Jr (May 1976) *ASTM Standardization News* 24–28