

The Role of Spectrophotometric Standards in the Clinical Chemistry Laboratory

Royden N. Rand

Department of Pathology, University of Pennsylvania, and William Pepper Laboratory, Hospital of the University of Pennsylvania, Philadelphia, Pa. 19104

(June 7, 1972)

It is obvious that erroneous data reported to a physician may adversely affect patient welfare. Currently, acceptable limits of accuracy and precision are poorly defined. It should be recognized, however, that the spectrophotometric measurement step in an appropriate analytical procedure is critical and inapparent error may occur. Spectrophotometric measurements, both manual and automated, are extensively used in the clinical chemistry laboratory. At least 1,000,000 such measurements per day on rather diverse equipment are made in this country; yet, few standards exist. Results of intra-lab surveys suggests that performance could be improved. The various ways in which spectrophotometry is used will be illustrated and a discussion of possible errors resulting from nonstandardized instrumentation will follow. There is pressing need for well defined and easily usable standards for wavelength, photometric accuracy, photometric linearity, stray light, as well as NBS specifications for optical cuvettes.

Key words: Clinical spectrophotometry, accuracy, precision; optical cuvettes; spectrophotometric standards, clinical.

I. Introduction

It is not widely appreciated that spectrophotometric methods represent the principal measurement techniques used in the clinical chemistry laboratory. A recent text lists 147 analytical techniques; of these, 84 require the use of spectrophotometry [1].¹ The twenty-five most widely used techniques at the William Pepper Laboratory are spectrophotometric.

The numbers of spectrophotometric analyses performed in any active hospital are impressive. At the Pepper Laboratory, for example, of 450,000 analyses per year about 300,000 require the measurement of the absorbance of light. We, therefore, make 800–1,000 spectrophotometric measurements per day. Other information allows us to estimate that more than 1,000,000 spectrophotometric tests are performed daily in the clinical laboratories in this country [2]. When it is realized that the growth-rate of clinical chemistry is approximately 15 percent per year, then a perspective of present requirements for accuracy and precision is immediately recognized.

Given the information that spectrophotometric measurements are so widely used, it seems strange that suitable standards are not readily available and in widespread use. It seems appropriate, therefore, for this communication to discuss:

1. The accuracy and precision requirements for clinical chemical measurements.
2. The use of spectrophotometry in the laboratory.
3. The present "state of the art" of spectrophotometric measurements.
4. The various types of spectrophotometric errors and how such errors may relate to the accuracy and precision requirements discussed in section III.
5. The type of standards which are required.

II. Nature of Clinical Laboratory Data

For the fields other than medicine, it is relatively easy to describe how the quantitative data of the analytical laboratory are used. Several examples can be cited:

1. The analysis of steel yields information that helps control quality, composition, and costs.
2. The analysis of gold is required to protect the purchaser's investment.
3. The analysis of the products of organic synthesis may verify composition and help optimize yield of desired products.
4. Quantitative analysis of biochemical systems helps elucidate reaction mechanisms.

¹ Figures in brackets indicate the literature references at the end of this paper.

In general, therefore, quantitative analysis can be used to obtain data that are used theoretically or empirically, in rather well defined ways.

By comparison, it is difficult to show clearly and unambiguously how quantitative clinical laboratory information is used in the practice of medicine. For example, while a change in the butterfat content of milk of 0.2 percent can be recognized as economically significant, a comparable change in the total protein content of serum, although easily measured, cannot be so easily interpreted. In general, small changes in the levels of medically important substances are difficult to interpret. Nevertheless, the clinical chemistry laboratory does quantitative analysis. It may be instructive, therefore, to describe how these data may be used.

A. Generalized Use In Diagnosis

Although recent studies of logic of diagnosis avoid dealing with the precise role of laboratory tests, it is generally recognized that they play a key role in medical decisionmaking [3].

When a physician sees a patient presenting certain signs and symptoms, he forms a hypothesis about the possible diseases compatible with them. He may then order one or more laboratory tests, the result of which will help confirm or negate the hypothesis. Given the likelihood of a disease, the expected increase or decrease in specific serum constituents helps confirm the presence of the disease. This use of laboratory data is essentially qualitative. That which is sought, is evidence of change in blood level clearly outside the normal range. Such use of laboratory information implies only moderate requirements of precision and accuracy.

From the viewpoint of the analyst, however, the perspective changes considerably. The concentration ranges of biologically important substances varies from micrograms to grams per deciliter. Further, the maximum concentration ranges encountered in disease may vary from 1.5 times the mean normal value to as much as 50 or greater. Because of this, the only possible means of assuring consistent information within and between laboratories, between analyst and between days, is to perform analyses with well defined and characterized methods, making measurements with the use of high purity standards and *instruments known to be calibrated correctly*.

B. Specific Diagnosis

To a limited extent, quantitative analysis provides specific diagnostic information. This is obviously true when the analytical procedure provides information concerning the amount of an abnormal substance—such as a paraprotein—present in a body fluid. There are, though, only a limited number of instances where this is the case, and indeed, few of these analyses use spectrophotometric procedures. Some substances normally in blood appear to vary within narrow limits. Examples of this are calcium and other electrolytes as well as total protein and albumin within an individual

[4, 5]. Further research in this area using measurements of high accuracy and precision may show how minimal abnormalities in levels of such substances are related to disease.

C. Therapy Related Decisions

An extremely important aspect of quantitative blood analysis is in relation to therapeutic decisions. The determination of such substances as sodium, potassium, hemoglobin, bilirubin, and certain drugs are commonly encountered examples. The accurate quantification of these substances seems to be required for consistent and reliable medical decision-making. For example, in fetal Rh incompatibility, pediatricians feel that a bilirubin level of 20 mg/dl suggests a high risk of significant brain damage and, thus, will intervene with exchange transfusion. Similarly, routine control of electrolyte levels following surgery requires impressive quantitative support from the laboratory. Another example is the quantification of the blood levels of therapeutic drugs to insure optimal response without toxicity. Accuracy and precision seem most clearly related to patient welfare for this application of quantitative analysis, where therapeutic decisions may depend on a blood level of some substance.

III. Accuracy and Precision Requirements

The discussion, to this point, suggests that accuracy and precision requirements of clinical laboratory determinations have not been rigorously defined. Several recent articles, however, have discussed the problem. The most comprehensive are by Barnett, entitled "Medical Significance of Laboratory Results," and Campbell and Owen, entitled "Clinical Laboratory Error in Perspective" [6, 7]. In both, the authors attempt to derive acceptable limits of variability for commonly used tests on the basis of the consensus of highly qualified physicians,² table 1 summarizes some of their findings.

The range of values for one standard deviation seems to be rather broad. For example, in Campbell and Owen, glucose was thought to require a reproducibility of 0.8–3.6 mg/dl; Barnett concludes that 5.0 mg/dl is acceptable. This variability may be due to the fact that the requirements for reproducibility were obtained by a consensus technique. Since there seems to be no current theoretical or experimental approach to the problem, consensus is the only means of arriving at these specifications.

Neither paper makes an explicit distinction between accuracy and precision. The medical decision levels represent the value at which a decision may be made and the variability represents the allowable precision limits. Although these are stated to be one standard deviation, it should be at least questioned as to whether these would not better represent two standard devia-

² Using operations research techniques in his doctoral dissertation, Cavanaugh has studied the probable effects of laboratory error for three commonly used tests [8]. He, too, found it necessary to use a consensus of informed medical expertise to elucidate the impact of laboratory error. He concludes that a "cost per laboratory error" can be computed and that this has meaning in terms "loss of life and limb." The concept is intriguing and Cavanaugh's challenging approach should be further explored.

TABLE 1. Acceptable Analytical Reproducibility

From Barnett [6] and Campbell and Owen [7]

Test	Decision level ¹ [6] mg/dl	Acceptable reproducibility ² [6] mg/dl	Acceptable reproducibility ² [7] mg/dl
Glucose	120	5.0	0.8 -3.6
Urea	27	2.0	0.9 -2.4
Calcium	11	0.25	0.03-0.33
Chloride ³	90	2.0	0.6 -2.7
Phosphate	4.5	0.25	0.14-0.25
Sodium ³	130	2.0	1.0 -2.6

¹ Level at which a medical decision may be made.² 1 S.D. in appropriate units.³ Units for chloride and sodium are milliequivalents per liter, mEq/l.

tion limits. In any case, the listed precision limits are certainly achievable as one standard deviation limit and probably for two standard deviation limits, by careful work.

For our purposes, it is not necessary to decide whether one figure or another is acceptable and correct. We may use them as guidelines to maximum acceptable error. In fact, we shall use Barnett's figures later in this paper to help elucidate some aspects of analytical error and, perhaps, better understand some of the requirements of spectrophotometric practice.

IV. Spectrophotometry in the Clinical Laboratory

A. Instrumentation

1. *Wide Band Instruments*—Wide band instruments are of many types and varieties with spectral bandwidths greater than 10 nm. These may be simple filter colorimeters or reasonably sophisticated spectrophotometers. Older instruments are null point type, whereas newer models are direct reading. Some are very simple; some are relatively complicated with highly stabilized electronics and digital absorbance readouts.

2. *Narrow Band*—The more expensive grating and prism ultraviolet visible type, narrow spectral bandwidth instruments are less commonly encountered. These may be single or double beam.

3. *Devices Used in Automation*—The most commonly encountered are double beam, interference filter, transmission type equipment.³

B. Type of Analyses Performed With Each Kind

1. *Wide Band*—The wide band instruments are primarily used for manually performed procedures in the visible area of the spectrum. In general, reagents are added to react with the substance being quantified

³ Strictly speaking these are not true double beam systems. One beam, of course, passes through the colored sample; the other passes through a filter of the same transmittance and is sensed by a separate photocell. Thus only variations in the output of the source can be compensated.

either directly in the fluid being analyzed or in a protein-free filtrate. A color is produced, sometimes after the application of heat, and its light absorbing characteristics are compared to those obtained with standards similarly treated.

2. *Narrow Band*—The most important applications of the narrow band instruments are: (1) the determination of enzymes, (2) toxicologically significant substances, (3) the direct absorptiometry of compounds such as bilirubin or uric acid, and (4) identification of unknown substances. In general, for these uses, absorptivities are used to compute the unknown concentrations.

3. *Photometric Instruments Used in Automation*—These instruments are used in a manner similar to the wide band instruments. They generally operate in the visible area of the spectrum and colors of standard compounds are compared to colors obtained with biological specimens.

V. "State of the Art" Spectrophotometric Measurement in the Clinical Laboratory

A. Wide Band Instruments

Few studies have been published on the significance of wide band instruments. A major study, however, was "Colorimeters—A Critical Assessment of Five Commercial Instruments" by Broughton and colleagues [9]. The participants studied two models of five manufacturer's instruments. The most significant performance factors were thought to be: (1) reproducibility (2) sensitivity (3) linearity. The overall design of the instrument, the accompanying sample cells and accessories were also critically discussed.

Stable solutions yielding absorbances of 0.05–0.55 were used to evaluate reproducibility. Typically, ten readings were made without resetting zero, then ten additional readings completed the series. Linearity was studied over a wide range of concentrations for five different chemical procedures, using five separate wavelengths. Sensitivity was defined as the ratio of the slope of the calibration curve to that obtained in a narrow band instrument.

A summary of their findings is found in table 2. Detailed examination of the paper makes clear that of the five wide band instruments examined, only one could be relied upon to yield calibration curves linear to an absorbance of 1, as well as sustaining a sensitivity in the range of 0.7 to 1. Although it is obviously unfair to extrapolate conclusions from this study to include all wide band instruments, it may be that many instruments of this type will exhibit some limitations on performance. It seems evident that the presence of such problems may limit the reliability of any analyses performed with such equipment. Standards for linearity and sensitivity evaluation are sorely needed.

Further evidence supporting this point may be found in a survey for wide band instruments performed in New York State [10]. Relevant data are found in

TABLE 2. Performance of Selected Wide Band Colorimeters

Reproducibility—when expressed as coefficient of variation, it varied from 1–3 percent at an absorbance of 0.1; above an absorbance of 0.4 it was 1 percent or less.

Linearity—varied from a continuous curve to linearity through an absorbance of one. Many instruments are linear only to absorbance of 0.5–0.6. Instruments of same make and type frequently showed significant differences in performance.

Sensitivity—widely variable, from a low of 21 percent to a high of 120 percent. Instruments of same make and type frequently showed significant differences.

Summary of data found in Broughton, et al. [9].

table 3. The three types of instruments listed are ones commonly found in clinical laboratories in this country. Three points are important:

1. The range of results reported for any solution, for any type of instrument is distressingly wide. For example, for instrument 1 and solution B, 95 percent of the instruments fall between 0.413 and 0.513.
2. There are significant deviations from linearity for instruments 2 and 3.
3. The relative sensitivity for instruments 2 and 3 are unimpressive.

The reasons for these poor results are not clear. Again, the need for appropriate standards and their frequent use seems very clear.

TABLE 3. Performance of Wide Band Spectrophotometers ¹

Instrument ²	Bandpass nm	Number of instruments	Absorbance		Deviation in linearity percent	Relative sensitivity percent
			Solution A ³ 0.0500M	Solution B ³ 0.100M		
No. 1	20	75	0.228 ± 0.012	0.463 ± 0.025	+ 1.5	95
No. 2						
10 mm	30	34	.191 ± 0.017	.382 ± 0.039	0	81
12 mm	35	58	.197 ± 0.011	.386 ± 0.022	– 1.8	82
19 mm	35	67	.193 ± 0.013	.368 ± 0.027	– 4.9	79
No. 3						
10 mm	20	33	.202 ± 0.022	.395 ± 0.044	– 2.5	84
12 mm	20	74	.214 ± 0.011	.416 ± 0.018	– 1.8	82
19 mm	20	51	.206 ± 0.008	.400 ± 0.019	– 3.2	85

¹ Vanderlinde, R., Instrumentation Survey—Visible Spectrophotometry. Report of Laboratories for Clinical Chemistry, New York State Dept. of Health, June 7, 1971.

² Millimeter designation refers to the size of the test tube used to hold sample.

³ Solutions A, B = 0.0500 M, 0.100 M cobaltous ammonium sulfate in 1 percent (v/v) H₂SO₄ read at 510 nm; all values corrected to 10 mm light path.

B. Narrow Band Instruments

It might be supposed that the performance of narrow band instruments would be more impressive than the wide band types. Unfortunately, this is a generalization not substantiated by published data. Several surveys bear this out:

1. Even if one looks only at the data for potassium nitrate, data from the world survey conducted by the Photoelectric Spectrophotometry Group [11], indicate a spread of results for both absorbance and peak wavelengths that is unacceptably broad.
2. In mid 1971, the College of American Pathologists (CAP) surveyed narrow band spectrophotometers in a number of clinical laboratories in the United States [12]. Solutions in sealed ampules included potassium dichromate (25, 50, 100 mg/l in 0.01 N H₂SO₄); Thompson Solution $\frac{1}{4}$, $\frac{1}{2}$ full strength); alkaline potassium chromate (40 mg/l in 0.05 N KOH). Data for some of these solutions are summarized in figure 1. The means and two standard deviations, as well as coeffi-

cients of variations are graphically shown. At 257 nm, the intra-instrument variability is aston-

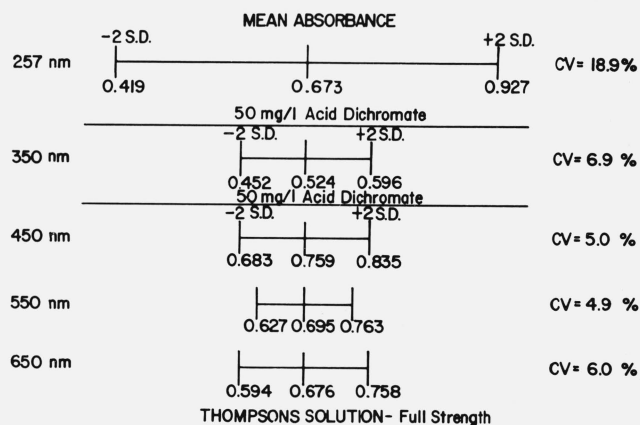


FIGURE 1. Performance of narrow band spectrophotometers in the United States.

Data from the report of the Subcommittee on Instrumentation, College of American Pathologists (1971).

ishingly high; at the other wavelengths, the coefficient of variation varied from 5-7 percent.

3. By contrast, a number of laboratories in New York State were surveyed at about the same time by the New York State Department of Health [10]. Some of these data are shown in figure 2. By comparison, the variability of instruments in the New York State survey was remarkably low.

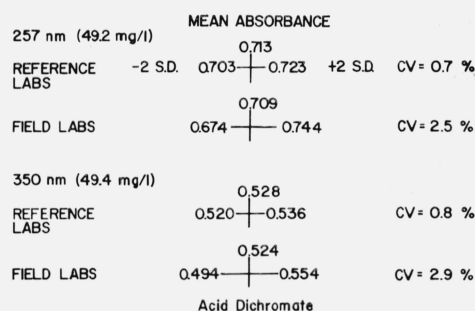


FIGURE 2. Performance of narrow band spectrophotometers in New York State.

Data from Vanderlinde, R., Report of a Survey, N.Y. State Department of Health (1971).

If the CAP results are representative of the range of spectrophotometric performance in this country, then we need to be concerned with the effect of such variable performance upon accuracy and precision. This point is amplified in table 4. The data represent the absorptivity of acid dichromate computed from the CAP survey for three concentrations. Shown are the high, low and mean values as well as data obtained in our laboratory. It is obvious that any analyses requiring an absorptivity for conversion of absorbance to concentration would be seriously in error, if a literature value were used.

It is also worth noting that the currently accepted value of 10.69 for the absorptivity of acid dichromate (in 0.1 H₂SO₄) was not met even at the mean [13]. Our values were also somewhat low. By contrast, we obtained 10.65 and 10.63 for the 50 mg/l and 100 mg/l solutions in the New York State survey.

TABLE 4. Absorptivity of Potassium Dichromate in 0.01 H₂SO₄

(350 nm)			
25 mg/l	High	12.24	
	Mean	10.48	
	Low	8.72	
	Pepper Lab	10.36	
50 mg/l	High	11.92	
	Mean	10.48	
	Low	8.04	
	Pepper Lab	10.58	
100 mg/l	High	11.83	
	Mean	10.09	
	Low	8.35	
	Pepper Lab	10.38	

¹ College of American Pathologists, Instrumentation Survey, May 1971 [12].

It is difficult to feel complacent about the reliability of measurements on narrow band instruments in the clinical labs in this country. When variability of ± 10 percent to ± 39 percent can occur, then surely the accuracy of analyses performed on such equipment must be severely questioned.

C. Automatic Instrumentation

To this author's knowledge, no studies of the photometers used in automatic instrumentation have been published. However, intra-laboratory surveys of analytical variability have been conducted [14]. Results have been analyzed by separation of the data from those laboratories using automatic methods from those using manual techniques. Typical data are shown in table 5. The reasons for the scatter in both instances include poor standardization and photometric variability. It would seem in order, therefore, to suggest that the photometric performance of automatic instrumentation should be studied in order to

TABLE 5. Results of an Intra-Laboratory Trial in Britain ¹

Substance ²	Manual methods				Autoanalyzer methods			
	Number	Mean concentration	S.D.	Range ³	Number	Mean concentration	S.D.	Range ³
Phosphorus	A. 131	3.037	0.406	1.2 to 9.3	35	2.860	0.273	2.3 to 3.7
	B. 135	7.758	1.231	3.1 to 10.4	35	7.906	0.576	6.6 to 9.0
	C. 134	4.458	0.469	3.1 to 5.5	35	4.666	0.243	4.0 to 5.3
Urea	A. 59	121.80	15.33	15 to 164	113	126.23	6.62	108 to 140
	B. 61	67.03	7.94	53 to 86	112	65.32	4.89	45 to 92
	C. 60	79.73	9.60	52 to 123	112	81.96	4.81	63 to 100

¹ From Gowenlock, A. H., Ann. Clin. Biochem. 6, 126 (1969).

² Samples A+B were dried sera; Sample C, aqueous solution.

³ Range includes reported results; mean and S.D. are best estimates.

assess the magnitude of the photometric component of analytic error. At the least, linearity specifications and standards seem to be required.

VI. Common Sources of Spectrophotometric Error and Their Possible Effect on the Accuracy and Precision of Clinical Chemical Analyses

A. Wide Band Instruments

1. *Wavelength*—Wavelength error is of considerable significance if colored substances with sharp absorption bands are analyzed. This is true, of course, because the rate of change of absorbance with wavelength will probably be high enough to cause a significant absorbance error. Usually, however, colored materials with appreciable broad absorption bands are analyzed in wide band instruments. Hence, clinical chemical analysis of this sort is little affected by wavelength error over moderate intervals, particularly if standards are run concurrently. On the other hand, it is well known that the optimum wavelength for any procedure is at the peak of absorbance. It follows that evidence of wavelength accuracy is implicit for reliable definition of this optimum wavelength.

2. *Photometric Accuracy*—It is difficult to give photometric accuracy unambiguous meaning in wide band instruments [13]. Thus, accurate absorbance measurements are not really possible in this type of instrument.

Some idea of the magnitude of the errors involved can be gained by study of table 6. This table summarizes an experiment in our laboratory in which a dilute solution of reduced nicotinamide adenine dinucleotide (NADH) was made in phosphate buffer at pH 7.4. Concentration 1 was twice concentration 2. These solutions were placed in 19 mm internal diameter test tubes and read against a phosphate blank in wide band grating instruments of 10 and 20 nm bandpasses. The same test tubes were used in both instruments and the readings were completed within a few minutes.

TABLE 6. Absorbance of NADH in Two Wide Band Instruments at 340 nm.^{1,2,3}

Instruments	Bandpass	Concentration 1	Concentration 2	ΔA
1	10 nm	0.304	0.156	0.148
2	20 nm	0.233	0.126	0.107

¹ NADH dissolved in 0.1 M phosphate buffer at pH 7.4.

² Concentration 2 was one-half concentration 1.

³ 10 mm round test tubes used for photometric readings.

Several conclusions are immediately obvious from these data: (1) Instrument 1 shows a more nearly linear relation of absorbance to concentration. (2)

Instrument 2 gives absorbances significantly below instrument 1. (3) If NADH coupled enzymes were being assayed and the NADH concentration changed from C_1 to C_2 , instrument 2 would have given an apparent enzyme activity 20 percent below instrument 1. Such differences, traceable to instrumentation, are most distressing. This is further evidence that the accuracy of spectrophotometric measurements in wide band instruments is questionable and may lead to significant error.

3. *Photometric linearity*—The wide band instruments used in most laboratories are assumed to be linear through an absorbance of 1.0. Earlier in this paper were cited data indicating variable performance of wide band instruments with respect to linearity [8]. It seems clear that the percentage error due to unsuspected nonlinearity is a function of the actual deviation and the absorbance at which nonlinearity becomes evident. Absorbance errors of more than 10 percent could be due to nonlinearity.

4. *Photometric Sensitivity*—Again, the British are to be credited for recognizing that linearity and photometric sensitivity are both important instrument parameters [9]. If an instrument is linear through an absorbance of one, then the slope of the line relating absorbance to concentration is a function of a number of variables. These include: (1) sample size, (2) absorptivity of colored complex, (3) optical path, (4) dilution factor, (5) photometric response. Since the slope and sensitivity may be altered by more than one factor, it is difficult to weight the photometric component and assign its contribution to analytical error. Indeed, there seems to be no general agreement as to what constitutes optimum sensitivity.

Some indication of the scope of the photometric problem may be obtained by reference to data found in table 7. This table lists for a few determinations the

TABLE 7. Required Sensitivity of Absorbance Measurements

Analyte	Medical decision level mg/dl.	Absorbance sensitivity required (Autoanalyzer)	Absorbance sensitivity required (Manual)
Glucose	120 \pm 5	0.013	0.005
BUN	27 \pm 2	0.010	0.080
Uric Acid	6 \pm 0.5	0.067	0.022
Calcium	11 \pm 0.25	0.035	0.080
Phosphate	4.5 \pm .025	0.016	0.080

From Barnett [6].

acceptable uncertainties pointed out by Barnett and shown in table 1 [6]. If we use actual absorbance-concentration relationships found for each of these, then for each of Barnett's uncertainties, we may compute a minimum change in absorbance which must be sensed and sensed repeatably. These are shown in

column 3 for Autoanalyzer⁴ procedures in our laboratory and in column 4 for manual procedures as described in a text [14]. Inspection of these data show the following; (1) The minimum sensitivity required is 0.005 absorbance and the maximum is 0.080. (2) Of the 10 values listed, five require a sensitivity of 0.022 or less. It seems likely that few current wide band instruments can consistently measure differences of 0.005; by the same token differences of 0.080 should easily be measured. It should also be realized that absorbance differences of 0.022 require, in general, measurement of transmittance values that differ by 1 percent or less. Thus, even to meet Barnett's broad tolerances requires spectral measurements of high sensitivity and precision.

B. Narrow Band Instrumentation

Narrow band spectrophotometers are used for the determination of a number of substances, some of which are listed in table 8. In addition to these, the

TABLE 8. Determinations Performed With the Use of Narrow Band Spectrophotometers

Bilirubin
NADH—coupled enzymes
Alkaline phosphatase (kinetic—using sodium dinitrophenol phosphate)
Barbiturates
Diphenylhydantoin
Doriden
Tolbutamide
Hemoglobin derivatives
Uric acid at 292 nm

partial characterization of the purity of standards by the determination of molar absorptivity is growing in importance. This use is discussed in this Journal by Burnett. (See figs. 1 and 2).

Our purpose in this section is to assess the significance of commonly encountered spectral errors and

their effect on the analytical reliability of the determinations involved. It would be beyond the scope of this article to exhaustively catalog each of the various determinations. Rather, we may take two different procedures and discuss some aspects of each.

First, let us consider the spectrophotometric determination of bilirubin [1]. Typically, a 1 to 50 dilution of serum may be made and the absorbance determined at two wavelengths, (455 and 575 nm). The 455 nm wavelength represents the bilirubin peak; the 575 nm peak may be used to correct the bilirubin absorbance for the contribution of hemoglobin. In table 9 some computations are based on two assumptions: (1) what we are dealing with is a hypothetical serum in which no interference from hemoglobin or other absorbing substances exist, (2) that the absorptivity of bilirubin in serum at 455 nm is 60.1.

If we refer to Barnett's paper we may note that the usual medical decision level for serum bilirubin in the newborn infant with hemolytic disease is 20 mg/dl and that the uncertainty of the bilirubin measurement at this level should be ± 1.5 mg/dl. One can, then, compute the expected absorbances that might be obtained at upper and lower limits; and this is shown in table 9, column 4.

It is to be noted that on the high side, the expected absorbance is 11 percent above the nominal and on the low side is 8 percent of the nominal. The "allowable" spread is in the range of 15–20 percent of the mean. In the previously cited CAP survey, the range of values reported for Thompson's solution at 450 nm, was 10 percent of the mean. Since one-half of the "allowable" variation could be accounted for by spectrophotometric error, only 5–10 percent remains for all other sources of error. It seems evident that reduction of the measurement error could significantly reduce analytical error. If this were so, the allowable limits might be set even lower than stated by Barnett.

Similarly, in table 10 some hypothetical computations with respect to NADH are made. That which is assumed is: (1) the "true" molar absorptivity of NADH is 6.2×10^3 , (2) a 10 percent high or low absorbance error

TABLE 9. Absorbance of Bilirubin Solutions and Medical Decisions¹

Substance	Decision level	Upper and lower limits ² mg/dl	Absorbance at limits	Absorbance difference	Percent of decision level
Bilirubin	20 mg/dl.	21.5	0.457	0.046	11
			.411		
		18.5	.379	.033	8

¹ Assumes a direct spectrophotometric procedure reading at 455 nm, an absorptivity of 60.1 and a dilution of 1/50 and no hemoglobin interference.

² See Barnett [6].

⁴ In order to adequately describe materials and experimental procedures, it was occasionally necessary to identify commercial products by manufacturer's name or label. In no instances does such identification imply endorsement by the National Bureau of Standards, nor does it imply that the particular product or equipment is necessarily the best available for that purpose.

can occur in any instrument, (3) the "true" absorptivity is used to compute enzyme activity in international units.

If such absorbance error occurs, then in column 3

TABLE 10. Spectrophotometric Error and Effective Molar Absorptivity of NADH and Its Effect on Computed NADH-Coupled Enzyme Levels

NADH level	Absorbance	Molar absorptivity	Apparent enzyme activity (U)	Enzyme activity (corrected for apparent absorptivity) (U)
Normal enzyme concentration			(Expected)	
7.8 mg/dl.	0.700	6.2×10^3	32.4	
10 percent low	0.630	7×10^3	32.4	28.6
10 percent high	0.770	5.7×10^3	32.4	35.0
Elevated enzyme concentration			(Expected)	
		6.2×10^3	404	
10 percent low		7×10^3	404	358
10 percent high		5.7×10^3	404	437

The high and low absorbance values are assumed as possible errors. See figure 1 for justification.

are listed the apparent molar absorptivities corresponding to the erroneous readings. In column 4 are listed activities for a hypothetical NADH coupled enzyme in the high normal range and a typical elevated enzyme activity. If we ignore the error occurring in the absorbance measurements of the enzyme activity itself, then in the last column are listed the "true" values assignable to these two enzyme activities based on NADH absorptivity corresponding to the high and low errors. The large differences are obvious.

Now this has the appearance of a quantitative argument; it is not intended to be so. It is, rather, a model to show the effect of spectrophotometric error on a clinically important determination. That such error can occur is attested to by the data in figure 1, from the CAP survey. Two coefficients of variation at 350 nm were observed to be 13.8 percent.

The common sources of spectral error in narrow band instruments are well understood. Wavelength error, nonlinearity, photometric inaccuracy, and stray light are major contributors. Standards for each of these are needed urgently.⁴

C. Automatic Instrumentation

Let us consider the photometers in use in automatic analyzers. The basic photometric requirements seem clear enough: (1) Given a chemistry which is linear, it is expected that the photometer output would exhibit basic conformance to the Beer-Bouguer Law over the range of transmittance of 0.1 to 1.0. (2) The sensitivity should be such that the desired or optimum absorbance concentration relationship is achieved and sustained from determination to determination.

In our hands, the various automatic analyzers seem to yield adequately linear calibration lines through an absorbance of one.⁵ Linearity should not be considered

apart from sensitivity since considerable flexibility exists with respect to possible ratios of sample to reagent. We may define the appropriate sensitivity as that which will yield a change in absorbance per defined detection limits (see table 2) clearly within the capability of the photometric system. This can be done by selection of the appropriate sample size; however, one may be hampered because the concentration of analyte which spans the 0-1 absorbance range may be so small as to require a large number of dilutions and re-running of samples. Some compromise is usually accepted.

An experiment was performed to evaluate the ability of an Autoanalyzer to discriminate small changes in concentration. Phosphate solutions in the range 3.23 to 4.73 mg/dl were prepared by dilutions of weighed reagent grade phosphate salt. These were analyzed on an Autoanalyzer (modified Sumner technique [17]), on three separate days. On each day, the standards used in the routine lab were utilized to prepare calibration curves. These were linear over the concentration range 1 mg/dl to 12 mg/dl (absorbance range 0.065 to 0.710). The equation of the standard line was computed by linear regression; this equation was then used to compute the values obtained for the weighed materials. Some of the data are summarized in table 11.

Over the three experimental days, the determined values group very closely with the maximum range not exceeding 0.22 mg/dl. Although all estimates are lower than the weighed-in values, the worst case underestimates by only 0.18 mg/dl. (See day 2-3.12 mg/dl.) Thus it would appear that at least for aqueous material, we can easily estimate phosphorus within 0.25 mg/dl. This falls within the specifications cited by Barnett [6].⁶

VII. Current Needs for Spectrophotometric Standards

The National Bureau of Standards has a current program related to spectrophotometric standards.

⁴ It is recognized that there are other sources of error which also contribute. Among these are variability in optical path dimensions, scatter of sample, fluorescence of sample, reflections, etc. Some of these were reviewed in a classic paper which indicates compelling obstacles to the determination of absolute absorbance [16]. Our argument here is that elimination of major sources of error would allow analytical accuracy at a level rarely approached. The question of the measurement of absolute absorbances and the elimination of all sources of error is a most difficult matter. The major effort now, however, should be to reduce within and between instrument variability and error to less than 1 percent.

⁵ This statement does not imply that there is evidence that all Autoanalyzer determinations are inherently linear. Some are; some are not.

⁶ It is possibly significant to mention that a skilled operator used the instrument. Furthermore, there is no evidence that the same sensitivity would apply to proteinaceous materials. To obtain data of this quality, standards must be run each time the analyzer is operated, at the very least.

TABLE 11. Ability of an Autoanalyzer to Distinguish Small Changes in Concentration Using Aqueous Phosphate Standards

Weighed concentration mg/dl.	Found mg/dl.			High-low range mg/dl.
	Run 1	Run 2	Run 3	
3.23	3.27	3.15	3.25	0.22
	3.27	3.05	3.16	
		3.15	3.25	
	Mean of 3 determinations = 3.19			
3.46	3.50	3.42	3.42	0.16
	3.43	3.51	3.42	
		3.42	3.34	
	Mean of 3 determinations = 3.43			
3.69	3.69	3.60	3.59	0.19
	3.69	3.60	3.50	
	Mean of 3 determinations = 3.61			
4.32	4.26	4.34	4.19	0.15
	4.26	4.34		
	Mean of 3 determinations = 4.29			
4.73	4.73	4.70	4.61	0.12
	4.73	4.61	4.61	
	Mean of 3 determinations = 4.66			

Current research has been summarized in this Journal and in another publication [18].

A major contribution of this program is the current availability of calibrated Schott NG Glass. These glasses, mounted in a convenient holder, are most useful for checking photometric accuracy in the visible range.

Other standards are also needed, if we are to hope for spectrophotometric measurements of uniformly high accuracy and precision. A summary of these follows:

A. Wavelength Standards

Suitable wavelength standards of universal applicability are strongly indicated. These should be usable in both narrow and wide band instruments.

B. Photometric Accuracy and Linearity

Although the current NG glasses are excellent, it appears that a more neutral glass, optically more homogeneous, would allow calibration to ± 0.1 percent relative transmittance. The capability of use in the ultraviolet would be a most important specification. Additionally, these glasses should be usable in wide band instruments.

Chemical standards are also needed for both narrow and wide band instruments for primary use in linearity and sensitivity evaluation. These would also be useful when necessary to prove that accurate spectrophotometric

measurements can be made with liquids. This in contrast, of course, to proof that the photometric accuracy of an instrument, itself, is acceptable. Chemical standards should also be available for checking the linearity and sensitivity of the photometers used in automatic analyzers.

Stray light, of major significance in ultraviolet measurements, needs a standard method for its measurement.

Spectrophotometric grade cuvettes need to be specified with respect to optical path length and wedge. A standard method of measurement needs to be established. Dr. Burnett's paper, I believe, indicates the importance of cuvette error in spectrophotometric systems.

Last, but not least, it should be emphasized that independent means for defining and measuring photometric accuracy need continued research. The presence of the high accuracy instrument at the Bureau should afford a useful tool for study of the problems related to this fundamental question.

VII. Summary

This paper has discussed the role of spectrophotometric standards in the clinical laboratory. Its underlying thesis is that errors in the color measuring step of photometric analysis have largely been ignored. Errors occurring in this step can and do contribute significantly to analytical error. It can be shown that errors traceable to the color measuring step can be of a magnitude such that medical decisions are made more difficult or may cause harm to the patient.

The assistance of Mrs. A. Ritz in obtaining some of the data in this paper is recognized with pleasure. The criticism of D. Arvan, who read the entire manuscript, was most helpful.

IX. References

- [1] Tietz, N. W., Ed., Fundamentals of Clinical Chemistry (W. B. Saunders Co., Philadelphia, Pa., 1970).
- [2] Brownfield, R. L., personal communication.
- [3] Feinstein, A. R., Ann. of Int. Med. **61**, 564 (1964).
- [4] Gordan, G. S., Loken, H. F., Blum, A., and Teal, G. S., Metabolism **11**, 94 (1962).
- [5] Cotlove, E., Harris, E. K. and Williams, G. Z., Clin. Chem. **16**, 1028 (1970).
- [6] Barnett, R. N., Amer. J. Clin. Path. **50**, 671 (1968).
- [7] Campbell, D. G. and Owen, J. A., Clin. Biochem. **1**, 3, (1967).
- [8] Cavanaugh, E. L., Doctoral Dissertation, Univ. of Calif., Berkeley, Calif. (1968).
- [9] Broughton, P. M. G., Riley, C., Cook, I. G. H., Sanders, P. G., Braunsberg, H., Colorimeters—A Critical Assessment of Five Instruments, Assoc. Clin. Biochemists (1966).
- [10] Vanderlinde, R., Report of A Survey, N.Y. State Dept. of Health (1971).
- [11] Anon., Photoelec. Spectrometry Group Bull. **16**, 441 (1965).
- [12] Anon., Report of the Subcommittee on Instrumentation, Coll. of Amer. Path. (1971).
- [13] Rand, R., Clin. Chem. **15**, 839 (1969).
- [14] Gowenlock, A. H., Ann. Clin. Biochem. **6**, 126 (1969).
- [15] Anon., Clinical Methods Manual, Bausch & Lomb, Inc., Rochester, N.Y. (1965).

- [16] Goldring, L. S., Hawes, R. C., Hare, G. H., Beckman, A. O. and Stickney, M. E., *Anal. Chem.* **25**, 869 (1953).
- [17] Sumner, J. B., *Science* **100**, 413 (1944).
- [18] Menis, O. and Shultz, J. I., Eds. *Nat. Bur. Stand. (U.S.) Tech. Note* 544, 151 pages (Sept. 1970), and 584, 175 pages (Dec. 1971).

(Paper 76A5-736)