

An Analytical Laboratory Experiment in Error Analysis: Repeated Determination of Glucose Using Commercial Glucometers

Paul L. Edmiston* and Theodore R. Williams

Department of Chemistry, The College of Wooster, Wooster, OH 44691; *pedmiston@acs.wooster.edu

Most introductory quantitative analysis courses introduce the concepts of statistical analysis as an important part of the evaluation of experimental data. To demonstrate the principles of statistics in the lab, students must carry out an experiment that can be performed in a short period of time so that many replicates are obtained yielding a statistically valid data set. Unfortunately, the relatively large amount of time required to complete most chemical analyses prevents the collection of a significant amount of data within the limits of a typical laboratory session; thus experiments used for a statistics lab are often simple in scope (1–4). For instance, Treptow reported an experiment that measured the volume of a graduated cylinder as a model for introducing precision and accuracy and to illustrate the difference between systematic and random errors in experimental results (4).

Although simple laboratory experiments such as these are instructive, they lack the real-world significance students tend to become enthusiastic about. To circumvent this problem, the first analysis experiment our students perform is the determination of glucose in aqueous solution using commercially available glucometers. Hundreds of thousands of diabetic individuals rely on the accuracy of these devices to control their blood glucose levels. The importance of the task of evaluating the glucometer's performance is easily conveyed to students, since in the United States diabetes is the fourth leading cause of death and the reliability of the devices is sometimes called into question. Recently, the Food and Drug Administration warned users of SureStep blood glucose meters manufactured by LifeScan Incorporated prior to August 1997 that the device may display an error message instead of "HI" when blood glucose concentrations exceed 500 mg/dL (5). Although evaluation of glucometer performance is found in the clinical care literature (6–9), this is the first report of its being done in an undergraduate chemistry laboratory.

What makes this laboratory attractive is that glucometer measurements can be performed in a matter of minutes with little practice by beginning analytical chemistry students, so that a large data set can be obtained rapidly. Students are required to calculate the range, mean, standard deviation, and percentage error for their measurements. They are then asked to pool their data and perform *F*-tests, *t*-tests, and *Q*-tests to either compare data sets or remove possible outlying data points. However, the most important part of the laboratory is to then have the students develop further experiments to assess the importance of a number of factors relating to the glucometer performance. These can include such things as determining the dynamic range or limit of detection of the meters.

Although it is also possible to use real blood samples instead of aqueous glucose solutions, this is problematic because of safety concerns involving possible pathogens in blood and because the glucose concentration in blood is unknown. Using aqueous solutions allows for accurate and safe test solutions. In addition, use of aqueous solutions allows students to consider the possibility of matrix effects, as there are many other compounds in blood that may be important to the measurement. If a blood matrix is required we suggest sterile bovine serum in order to avoid problems with pathogens.

Principle

The glucometer used in this study is the Encore model available from Bayer Corporation, Diagnostic Division, Elkhart, Indiana. Samples are applied to disposable test strips that consist of a membrane which immobilizes reagents necessary for a series of enzyme-catalyzed reactions. Hexokinase catalyses the reaction of adenosine triphosphate and magnesium ion with glucose in the sample to produce glucose-6-phosphate. The glucose-6-phosphate is substrate for glucose-6-phosphate dehydrogenase, which converts NAD present in the strip to NADH. Next, diaphorase catalyses the reaction between NADH and tetrazolium to produce the brown compound formazan, which is directly proportional to the plasma glucose concentration in the whole blood sample. The reflectance of light from the test strip membrane that contains the reacted materials is measured to quantify the glucose concentration. Detection is provided by a simple photocell. There is a direct relationship between the quantity of reflected light and the amount of glucose present in the whole blood sample. The test strips are one-time use only, are disposed of after each individual measurement, and cost about 50–90¢ each depending on the model of the glucometer. Glucometers cost approximately \$50.

Experiment

In the initial experiment each student is assigned a specific glucometer to be used for all measurements. Glucose standards and unknown aqueous samples are provided for each student. Students are also given a sufficient number of test strips to complete the assignment. Each student then measures both a glucose standard and an unknown six times. All results reported are required to include means, standard deviations, and confidence limits for the samples. The class data are compiled and analyzed as a whole, and relevant statistical tests are performed. The experiment requires the students to use

t , F , and Q tests to evaluate their data within the framework of all the data obtained by the class. Below you will find examples of typical student results.

After evaluation of the data obtained from the assigned glucometers, students are asked to design their own experiments to answer one of several questions that may arise after analysis of the initial data set. These questions have included:

1. Do outdated test strips give the same results as new strips?
2. Does the time for development of the color matter?
3. Does the volume of the sample influence the results?
4. What is the dynamic range of the glucometer?
5. Are the meters equally precise throughout the entire dynamic range?
6. Do different meters give the same result on the same sample and why?
7. How are samples prepared? Does the composition of the sample affect results?
8. How is the glucometer calibrated?

In this manner the student must design and carry out his or her own experiment, which provides an additional opportunity for learning. In addition, important introductory topics in analytical chemistry such as sampling, sample storage, instrument reliability, dynamic range, matrix effects, and calibration can be framed within this experimental setting. The lab can thus be easily integrated into lecture topics on statistics using student experiments as relevant examples.

Discussion

After the students have completed their work it is invaluable, for three reasons, to compile all the results and present them for discussion to the class. First, because all the data obtained by the students are examined at one time, measurements that appear faulty or invalid are often easily discernible. Second, data sets can be compared to answer important questions. For instance, do different glucometers give the same value statistically for the same standard sample? Third, if inconsistencies such as two glucometers giving different results for the same sample are found, experiments can be designed with the oversight of the instructor to determine the cause of inconsistencies. To help illustrate these points examples from data obtained from students performing this laboratory are presented.

Case 1: Outliers

It has been our experience that within a class of 20–30 students there are often several glaring outliers, which allow for instruction on how data can be removed using the Q -test. Poor measurements such as these are most often caused by improper application of sample to the test strips. A complete data set demonstrates how important replicates are in quantitative chemical analysis. A highly inaccurate low measurement taken without replicates could preclude a diabetic's administering insulin when the actual glucose level is several times higher than the anomalous measurement indicates; this illustrates that a decision based on poor data has real consequences.

Case 2: Comparison of Two Data Sets

In many instances there are statistically valid differences between students' data sets. For example, Table 1 lists data taken by two students performing the same analysis on the same aqueous glucose sample using two different glucometers. Taken individually, each set of data is very representative of the precision obtained by students using the Encore glucometers.

Examining the mean values obtained by each student it appears that student B, using glucometer #2, has measured glucose concentration values that appear to be statistically higher than those of student A. Performing a t -test allows this to be proven. Using a calculated s_{pooled} of 5.06, one calculates t to be 4.85, which indicates these data sets are statistically different even at the 99% confidence level.

After determining that these data sets are indeed different, the obvious question in students' minds is which set of data is more accurate? Until this point in the discussion, which takes place after the experiment has concluded, the concentration of the glucose standards is not revealed. In general this is done to prevent the data from being biased and to give the instructor flexibility in the learning process. For the case presented here, the standard sample being measured was a pure aqueous solution of glucose with a concentration of 150 mg/dL. Therefore, it is evident that both students have readings that are low compared to the "actual" value. (In this case this difference may be due to a matrix effect, since the standard is a pure aqueous glucose solution whereas the meter is designed for use with whole blood.) To determine if there are statistical differences from the "actual" concentration, one-population t -tests can again be performed in comparison to the nominal value. For both sets of data there is a statistical difference from the nominal value even at the 99% confidence level.

F -tests can also be introduced. It should be noted that the variances between Encore glucometers are typically not different (this includes the data presented here). However, variances between different brands of glucometers have been found to be different; thus two-population t -tests can be utilized in this laboratory experiment as well.

Case 3: Method Validation

Since a large amount of quantitative data can be obtained quickly, this lab can be used to introduce concepts such as calibration curves, limits of detection, and dynamic range.

Table 1. Results on the Same Aqueous Glucose Standard Using Different Glucometers

Trial	Measured Glucose Concn/mg dL ⁻¹	
	Student A Glucometer #1	Student B Glucometer #2
1	121	128
2	127	133
3	126	141
4	118	144
5	120	140
6	121	132
mean	122.2	136.3
s	3.5	6.2

As an independent experiment, students can be assigned to make up their own glucose solutions in order to test the analytical performance of the glucometers.

Case 4: Developing Experiments

In addition to method validation, the initial test data obtained by the class will pose many questions about precision and accuracy of the glucometers. For instance, in the last scenario where sets of data were compared, two obvious questions arose: (i) Why do the two glucometers give different results? and (ii) Why is there a consistent deviation from the "actual" value of the glucose standard? To answer these questions students are asked to develop hypotheses and experiments to test them. Students can even be asked to submit a formal proposal if desired. For the question of why the glucometer readings are consistently low, students concluded that either there was a systematic error in the measurement or the indicated concentration of standard was incorrect. Thinking that the inconsistencies were due to systematic error, one student hypothesized that because the measurement is based on reflectance, a colorless aqueous sample would not be equivalent in color to blood, which is the matrix the glucometer was designed to test. The student proposed repeating the measurement but adding a dye to the aqueous standard to mimic the color of the blood. Other students hypothesized that the standard sample was degraded by microorganisms during storage prior to the laboratory and tested the hypothesis that cold-temperature storage of solutions would lead to more accurate results. The amount of in-depth analysis and experimentation can be great, so the laboratory can be extended over a multi-week period. During this time diverse topics such as dynamic range, clinical chemistry, and possible development of implantable glucose sensors can be discussed, depending on the scope of the course.

Summary

Reported here is an interesting yet time-economic experiment using commercially available glucometers to make

replicate measurements for the demonstration of statistical analysis in the analytical laboratory. Often results can lead to in-depth questions and further experiments which focus on error analysis and the reliability of the glucometers. The measurement is both stimulating and relevant to the students. The major drawback in this experiment is the expense of the meters and test strips; however, the meters can easily be shared by two or three students and waste disposal is minimal.

Acknowledgment

P. L. Edmiston was supported as a Camille and Henry Dreyfus Fellow at The College of Wooster under the mentorship of Charles L. Borders Jr.

^WSupplemental Material

Supplemental material for this article is available in this issue of *JCE Online*.

Literature Cited

1. Sen, B. *J. Chem. Educ.* **1977**, *54*, 468.
2. Suder, R. *J. Chem. Educ.* **1989**, *66*, 437.
3. Guedens, W. J.; Yperman, J.; Mullens, J. *J. Chem. Educ.* **1993**, *70*, 776–779. Guedens, W. J.; Yperman, J.; Mullens, J. *J. Chem. Educ.* **1993**, *70*, 838–841.
4. Treptow, R. S. *J. Chem. Educ.* **1998**, *75*, 992–995.
5. United States Department of Health and Human Services, Food and Drug Administration. Press Release *P98-20*, July 28, 1998.
6. Harrison, B.; Markes, R.; Bradley, P.; Ismail, I. A. *Clin. Biochem.* **1996**, *29*, 521–527.
7. MacKinnon, D. T.; Henderson, A. R. *Clin. Biochem.* **1994**, *27*, 501–505.
8. Devreese, K.; Leroux-Roels, G. *Eur. J. Chem. Clin. Biochem.* **1993**, *31*, 829–837.
9. Usmani, H. A.; Khan, I. I.; Mughal, F. H. *JPMA J. Pak. Med. Assoc.* **1998**, *48*, 114–116.