Performance Requirements for Glucose Assays in Intensive Care Units

James C. Boyd1* and David E. Bruns1

Treatment of patients in the intensive care unit (ICU)2 with intravenous insulin became popular after initial reports suggesting that such treatment led to decreased in-hospital and 30-day mortality (1–3). Such treatment became controversial when several subsequent studies failed to confirm the improved clinical outcomes (4–6). The latter studies, however, used different methods for glucose measurement, and the measurements typically were made less frequently (7).

If we wish to optimize glucose monitoring in patients being treated with intravenous insulin while in the intensive care unit, we need answers to several questions: 1) How frequently should glucose be measured? 2) What quality of glucose measurement is needed? 3) What protocol should be used to guide insulin treatment? 4) How much nursing time will be required to conduct such monitoring? 5) Will such monitoring lead to improved outcomes? This editorial concerns itself primarily with the first 3 of these, but the last 2 are extremely important considerations in the clinical arena.

Computer simulation studies offer several advantages for answering such questions. First, they entail no risk to patients. Second, many thousands (or millions) of patient-care episodes can be simulated; moreover, various treatment protocols and the effects of different measurement frequencies can be considered simultaneously with evaluation of the analytical requirements for glucose assays. Finally, simulations can predict surrogate markers of patient outcomes such as rates of hypoglycemia and hyperglycemia, time in range, and within-patient variation of glycemia.

This issue of Clinical Chemistry contains 2 reports of novel computer simulation studies of glycemic control in ICU patients (8, 9). These studies may help to answer questions regarding what analytical quality is required of glucose assays, and what impact more-frequent measurement has on patient outcomes and/or the analytical performance requirements of glucose measurement. Although earlier modeling studies of these questions have used established physiologic models of glucose homeostasis (10, 11), what makes the current modeling studies unique and so important is that the models in these studies are based on data from actual ICU patients.

Modeling

More than 350 models of glucose homeostasis were published as of 2003; the number is now undoubtedly much larger. In our simulation studies (10, 11), we chose a model based on the classic minimal model of Tofello et al. (12). Although the minimal model, in its many variations, has drawbacks, the approach has been deemed “indisputably useful” (13). This method offers the advantage of freely manipulable parameters allowing the exploration of a wide range of starting glucose concentrations, a wide range of constant or varying insulin sensitivities, and a long duration of monitoring (and thus stable estimates of glycemic variability, etc.), as well as a range of realistic values for analytical imprecision and inaccuracy. On the other hand, the model is difficult to validate as being representative of a given patient population because glucose fluctuations in patients are influenced by many factors unique to the physiology of each individual and/or the method used for glucose monitoring in a given patient. One of the most difficult problems in performing computer simulation studies is how to demonstrate that the results obtained from simulated patients actually translate to the care of real patients.

The 2 studies in the current issue of Clinical Chemistry (8, 9) take alternative and differing modeling approaches on the basis of real patient data. The modeling performed by Wilinska and Hovorka (8) was performed in a population of 56 “virtual” patients, where each patient was modeled individually by fitting the data from an actual patient episode in the ICU to a complex, physiologically based compartmental model that uses differential equations. Their model was validated by demonstrating that it could replicate the principal findings of 2 open-label randomized clinical trials of glucose control protocols. Wilinska and Hovorka...
considered 765 different combinations of proportional bias, constant bias, residual error, and autocorrelation in their modeling scheme. Owing to the computational complexity of their model, they considered only data from a random sampling of 60 of the 765 combinations (20 combinations for each of the 3 ranges of measurement error that they considered) where the data were simulated for a 48-h ICU stay. They compared 3 different glucose-control protocols in their simulations and simulated data for continuous glucose monitoring (CGM) and intermittent glucose monitoring.

Van Herpe et al. (9) took a different approach. These investigators recorded the interventions actually taken in ICU patients on the basis of glucose measurements made with blood gas analyzers, and compared those interventions to the ones that would have been taken when using less-accurate simulated glucose sensors. Each new (simulated) intervention was compared with the original (nominal) intervention by use of a 3-dimensional error grid system, and the differences were graded into 4 categories (A, B, C, D) of increasing error severity, with error limit thresholds established by an expert panel. These investigators calculated the probability of occurrence of each error type for glucose sensors with varying levels of bias and imprecision for both simulated CGM and simulated intermittent glucose monitoring scenarios.

Although the goal of each of these studies was the same—to provide advice regarding the performance requirements for CGM sensors on the basis of simulation modeling—the simulation results obtained by the 2 differing approaches were evaluated by use of different outcome measures. The outcomes documented by Wilinska and Hovorka (8) were more traditional, based on rates of hypoglycemia, hyperglycemia, time in range, and glycemic variability, whereas those in the study of Van Herpe et al. (9) were the probabilities of an error occurring in each error category. Despite these differences in outcome measures monitored, results of both studies suggested that CGM glucose sensors with mean absolute relative difference (MARD) scores <11% gave the best results and lowest frequencies of hypoglycemia.

MARD scores may be unfamiliar to many readers of Clinical Chemistry. MARD scores are calculated on the basis of the mean percentage difference of the absolute concentration measured by the evaluated glucose sensor from a reference glucose concentration. MARD is widely used by manufacturers of CGM devices for expression of the overall assay error. MARD scores are related (but not in a simple way) to the scores for total analytical error (TE) that are more familiar to clinical laboratorians (TE = bias [%] + z × CV [%]), where z is a factor used to define the probability that results exceed the TE threshold. Assays with bias equal to +10% or −10% but zero imprecision will have an identical MARD score of 10% with no indication of the direction of difference. However, for assays with zero bias, a MARD score of 10% will correspond to an assay with an imprecision (CV) of 12.5% (TE of ±20.6% when z = 1.65) on the basis of the theoretical relationship of CV = √(π/2) × MARD for normally distributed data (14). Defining the relationship between MARD and TE becomes more difficult with non–normally distributed data or when both the assay bias and imprecision are nonzero. Based on MARDs calculated from sensors of a given TE, Van Herpe et al. have expressed this relationship graphically in Fig. 4 of their paper (9).

The optimal MARD scores of <10 or <11 suggested, respectively, in Wilinska and Hovorka (8) and Van Herpe et al. (9) for CGM, in which glucose is measured frequently, correspond to upper TE limits as large as 20%–23% when z = 1.65. Such a TE allowance is much larger than the TE limits currently under consideration for intermittent glucose measurement systems. The findings from these 2 studies suggest that the acquisition of glucose data more frequently by a CGM system carries definite advantages for management of glycemia, even when the analytical performance of a CGM is inferior to that recommended for meters used in intermittent glucose monitoring. The results in these 2 reports also support the conclusion of our earlier studies, which used a third simulation modeling approach: quality specifications for imprecision of glucose meters are not transferable to CGM (11).

The models agree that control of glycemia is intimately related to (a) the analytical performance of the glucose measurement system, (b) the frequency of the measurements, and (c) the protocol by which a measured glucose is translated into an intervention. Moreover, these 3 determinants are interrelated. The implications for design of glucose-measuring systems are becoming clearer, as are implications regarding the frequency of measurements in the face of constraints on nursing time in ICUs. Given the growing evidence that patient outcomes, including mortality, are predicted by control of glycemia, these issues demand attention.

All modeling studies carry the same conundrum expressed years ago by the statistician George E.P. Box, who stated, “Essentially, all models are wrong, but some are useful” (15). All of the models described to study the question of the effects of measurement imprecision on control of glycemia are admittedly imperfect. But the fact that 3 independent models, distinct from each other in their construction, produce similar conclusions makes the results of those models more credible. Wilinska and Hovorka and Van Herpe et al. are to be congratulated for the valuable contributions their modeling studies have made in suggesting poten-
tial new criteria for CGM performance that not only are in agreement, but are based on real patient data.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the author disclosure form.

Disclosures and/or potential conflicts of interest:

Employment or Leadership: J.C. Boyd, Clinical Chemistry, AACC.

Consultant or Advisory Role: None declared.

Stock Ownership: None declared.

Honoraria: None declared.


Expert Testimony: None declared.

Patents: None declared.

References