Clinical Chemistry 61:4
666–670 (2015)

Letters to the Editor

Using the Wrong Model Can Lead to Unsupported Conclusions about Glucose Meters

To the Editor:

Van Herpe et al. *(1)* modeled the effect of glucose sensor errors to provide total error acceptability limits to prevent harm to patients. In particular, the authors state that if the total error is <15.7%, the probability is zero that glucose meter results will fall in the D zone (causes severe injury or death) of a glucose meter error grid.

This cannot be true for several reasons. The authors simulate total error by sampling from a gaussian distribution. They may well have observed zero results in the D zone, but this is not the same as claiming a zero probability of D zone results. The gaussian distribution ranges from minus infinity to plus infinity, so as long as the SD is not zero, it is a mathematical certainty that the probability of results larger than 15.7% is greater than zero.

But perhaps more important is the incompleteness of the error model chosen by the authors. They have not modeled the effects of interferences, which have previously been shown to contribute to total error and are independent from average bias and imprecision *(2)*. Granted that interferences are difficult to model, but a survey has shown that they are a significant source of clinician complaints about laboratory error *(3)* and have caused injury and death related to glucose meter use *(4)*.

Additionally, one might infer from the authors' results that specific combinations of imprecision and bias will provide acceptable results, but Krouwer *(5)* has shown that failing to include interferences in the model can be misleading, especially for glucose meters, where interferences are common and increase the total error beyond that modeled for bias and imprecision.

Van Herpe et al. state that user errors, as well as several other effects, have been omitted from their model. Whereas some types of user error will affect results regardless of the meter, harm to patients nevertheless occurs.

There is always a risk of D zone errors. Risk analysis with methods such as failure mode effects analysis and fault tree analysis are an effective way to minimize the risk of large, rare errors.

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: *No authors declared any potential conflicts of interest.*

References

1. Van Herpe T, De Moor B, Van den Berghe G, Mesotten D. Modeling of effect of glucose sensor errors on insulin dosage and glucose bolus computed by LOGIC-insulin. Clin Chem 2014;60:1510–8.
2. Lawton WH, Sylvester EA, Young-Ferraro BJ. Statistical comparison of multiple analytic procedures: application to clinical chemistry. Technometrics 1979;21:397–409.
3. Krouwer JS. Estimating total analytical error and its sources: techniques to improve method evaluation. Arch Pathol Lab Med 1992;116:726–31.
4. US Food and Drug Administration. Advice for patients: serious errors with certain blood glucose monitoring test strips. http://www.fda.gov/MedicalDevices/Safety/AlertsandNotices/PatientAlerts/ucm177189.htm (Accessed December 2014).
5. Krouwer JS. The danger of using total error models to compare glucose meter performance. J Diabet Sci Technol 2014;8:419–21.

Jan S. Krouwer*

Krouwer Consulting
Sherborn, MA

* Address correspondence to the author at:
Krouwer Consulting
26 Parks Dr
Sherborn, MA 01770
Fax 508-653-2379
E-mail jan.krouwer@comcast.net

Previously published online at
DOI: 10.1373/clinchem.2014.237065

In Reply

We thank Dr. Krouwer for his constructive suggestions regarding the total error modeling strategy as recently presented in Van Herpe et al. *(1)*. We agree that, from a mathematical point of view, the probability of glucose meter results falling in the D zone should not be zero for a real meter with a total error below the accuracy threshold. As clearly stated in our work, we formulated this conclusion only in the scope of the executed simulations. Although the number of simulations was high compared to alternative simulation studies in this field, it does not imply a 100% guarantee for future real behavior. Risk analysis techniques (such as failure mode effects analysis) are complementary to simulation studies and are even essential in the regulatory process (CE marking in Europe, Food and Drug Administration approval in the US) of such medical devices, but fell outside the scope of the work.

We also agree with Dr. Krouwer that the error mode used, though forming the base of similar simulation studies *(2–4)*, is incomplete, as random patient interferences were not included. Together with the pre- and postanalytical errors as mentioned in our article *(1)*, these are factors that will undoubtedly increase the real total error. This is exactly the reason we advised

in our work to adopt accuracy requirements more stringent than those resulting from simulations.

Simulations are always based on assumptions, and unfortunately, a model is only a model. An extended version of the error model will remain another approach of reality. What is probably more important to reduce errors when defining clinically realistic accuracy thresholds is that our study, to the best of our knowledge, is the first that is based on glucose dynamics originating from real-life critically ill patients (i.e., independent of any mathematical glucoregulatory model and avoiding the associated errors). Further, clinical studies to validate a (new) glucose sensor should be appropriately designed (sufficient number of target patients, adequate reference sensor, etc.) to compare its accuracy performance to such thresholds. Next, glucose sensor accuracy thresholds do depend on the robustness of the control algorithm and should be specified (using simulations) for each individual glucose controller *(1, 5)*. Generalization of these thresholds will underestimate errors for less robust glucose controllers and potentially harm patients; it should be avoided, accordingly. Finally, we wish to underline the need for clinical trials investigating the combination glucose sensor/glucose controller (each with its specific characteristics) in a real-life critically ill setting to overcome the shortcomings typical of simulation studies.

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: None declared.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: Supported by the Agency for Innovation by Science & Technology IWT-TBM program (100793), KU Leuven IOF-HB-10/039, and KU Leuven IOF-HB-13/027. B. De Moor, GOA/10/09 and iMinds Medical Information Technologies SBO 2014; G. Van den Berghe, Methusalem program of Flemish government (METH08/07) and ERC Advanced grant (AdvG-2012-321670); D. Mesotten, Senior Clinical Fellowship from Research Foundation–Flanders.
Expert Testimony: None declared.
Patents: None declared.

References

1. Van Herpe T, De Moor B, Van den Berghe G, Mesotten D. Modeling of effect of glucose sensor errors on insulin dosage and glucose bolus computed by LOGIC-insulin. Clin Chem 2014;60:1510–8.
2. Karon BS, Boyd JC, Klee GG. Glucose meter performance criteria for tight glycemic control estimated by simulation modeling. Clin Chem 2010;56:1091–7.
3. Boyd JC, Bruns DE. Quality specifications for glucose meters: assessment by simulation modeling of errors in insulin dose. Clin Chem 2001;47:209–14.
4. Boyd JC, Bruns DE. Effects of measurement frequency on analytical quality required for glucose measurements in intensive care units: assessments by simulation models. Clin Chem 2014;60:644–50.
5. Wilinska ME, Hovorka R. Glucose control in the intensive care unit by use of continuous glucose monitoring: what level of measurement error is acceptable? Clin Chem 2014;60:1500–9.

Tom Van Herpe[1,2]*
Bart De Moor[2]
Greet Van den Berghe[1]
Dieter Mesotten[1]

[1] *Department of Intensive Care Medicine*
Katholieke Universiteit Leuven
University Hospitals Leuven
Leuven, Belgium
[2] *Department of Electrical Engineering (ESAT)*
Katholieke Universiteit Leuven
Research Division SCD
iMINDS Medical Information Technologies
Leuven (Heverlee), Belgium

* Address correspondence to this author at:
Department of Intensive Care Medicine
KU Leuven–University Hospitals
Leuven
Herestraat 49
B-3000 Leuven, Belgium
Fax +32-16-344015
E-mail tom.vanherpe@esat.kuleuven.be

Previously published online at
DOI: 10.1373/clinchem.2014.237420

Is Ferrotoxicity a New Great Public Health Challenge?

To the Editor:

The recent report of ferrotoxicity as a marker of increased risk of mortality by Ellervik et al. *(1)* needs to be viewed against the background of the study design, its weaknesses and strengths. Likewise, the metaanalysis in that study, which the authors found supportive of their own findings, needs to be judged against the fact that, of 72 relevant studies identified by the search strings they used, they threw out 70; thus, only 2 other studies besides their own entered into the metaanalysis. Their main finding, that high serum ferritin is associated with increased mortality in this cohort, may simply be because ferritin is an acute-phase reactant, and individuals affected by various chronic diseases often have a chronic inflammatory state that includes raised ferritin as part of the inflammatory biomarker signature. Although the authors adjusted for well-recognized major risk factors (modifiable and unmodifiable), no adjustments for inflammatory status appear to have been made *(1)*, although it is known from previous publications that at least C-reactive protein is available for this cohort. It would be interesting to learn why the authors chose not to adjust for this, because it seems to be a flaw in the