ORIGINAL ARTICLE

Technical Aspects of the Parkes Error Grid

Andreas Pfützner, M.D., Ph.D.,[1] David C. Klonoff, M.D.,[2] Scott Pardo, Ph.D.,[3]
and Joan L. Parkes, Ph.D.[3]

Abstract

Background:

The Parkes error grid, which was developed in 1994, presented performance zones for blood glucose (BG) monitors with borders that were not mathematically specified at the time the grid was published.

Methods:

In this article, we (1) review the history of the Parkes error grid, (2) present the never-before-published exact coordinates and specifications of the grid so that others may produce an exact replica of the original grid, and (3) discuss our suggestions how this metric should be applied.

Results:

The new ISO15197:2013 guideline for system accuracy assessment of BG meters for patient self-measurement incorporates use of this metric for defining acceptable accuracy of BG monitors. It is expected that, for regulatory purposes, this document will stipulate that the error grid version for type 1 diabetes should be applied with the caveat that only the A zone represents acceptable accuracy.

Conclusions:

It remains to be seen by how much the new error grid, which is currently being developed by the Food and Drug Administration/Diabetes Technology Society/American Diabetes Association/The Endocrine Society/Association for Advancement of Medical Instrumentation, will deviate from the Parkers error grid.

J Diabetes Sci Technol 2013;7(5):1275–1281

Introduction

The Parkes error grid was published in 2000 based on a survey of 100 physician attendees at the June 1994 American Diabetes Meeting.[1] This metric has been intended for use in assessing clinical accuracy of blood glucose (BG) meters for patient self-measurement. By developing this new grid, the authors intended to provide an alternative to the original Clarke error grid,[2] which was originally developed as a teaching tool rather than a BG meter clinical accuracy assessment tool and had been criticized for the placement of its risk boundaries.[3,4] The clinical accuracy of a BG value

measured by a BG monitor can be expressed as a description of the potential clinical outcome associated with basing a treatment decision on this value. Clinical accuracy is focusing on the clinical relevance of the meter results in comparison with analytical accuracy, which gives detailed information about agreement of the home BG meters in comparison with a reference method. The Parkes error grid has been introduced as an accepted evaluation tool in the new draft ISO15197:2013 guideline.[5]

The Parkes error grid was developed in June 1994, which was less than 1 year after publication of the Diabetes Control and Complications Trial study. At that time, intensive insulin therapy was not widely accepted, analog insulins were not available, and acceptable analytical accuracy (for BG readings of >75 mg/dl) was ±20%.[6,7] This level of accuracy has been replaced by more rigorous requirements in the new ISO15197 standard. For these and other reasons, there is a consensus in the clinical chemistry, regulatory, and medical community that a modern error grid is needed. For this reason, an expert panel (organized by the Food and Drug Administration, the Diabetes Technology Society, the American Diabetes Association, The Endocrine Society, and the Association for Advancement of Medical Instrumentation) is currently developing a next-generation error grid to describe the clinical performance of BG monitors by incorporating modern treatments and performance expectations.

The Parkes error grid (like the Clarke error grid) specifies five risk levels. The article by Parkes and coauthors,[1] which presented the Parkes error grids (including one for type 1 diabetes patients and one for type 2 diabetes patients on insulin therapy), depicted boundaries for the performance zones but did not present technical specifications or coordinates.[1] Such information would be necessary to allow others to generate their own mathematically accurate versions of these grids and performance zones. It has therefore been necessary to scan or trace the Parkes error grid boundaries, some of which are curved, in order to work with an exact version of this metric.

The purpose of this article is to provide previously unpublished technical information (i.e., coordinates for constructing the Parkes error grid and how the grid was constructed) about the grid and to discuss interpretations of using the Parkes error grid in the evaluation of BG meter systems for medical, regulatory and marketing purposes.

Creation of the Parkes Error Grid

Parkes error grids have been separately defined for two types of diabetes: persons with type 1 diabetes and persons with type 2 diabetes on insulin therapy. This differentiation reflected the opinion of the Parkes survey respondents that patients with type 2 diabetes on insulin can potentially tolerate a larger error in the BG meter value than can patients with type 1 diabetes and led to development of two grids. The type 2 diabetes error grid called for less rigorous clinical performance than the type 1 diabetes grid called for, especially in the low glucose range, and it has fallen out of favor. Many clinical chemists are not even aware that the widely used Parkes error grid was originally intended only for type 1 diabetes patients and not for type 2 diabetes patients on insulin. The risk definitions and the boundaries or coordinates (which have not been previously published) are provided in **Table 1** and **Figure 1**.

The Parkes error grids were created as follows: 100 physician attendees of the 1994 American Diabetes Association meeting completed two separate surveys, one for type 1 diabetes patients and the other for type 2 diabetes patients, first by defining five BG ranges as defined in **Table 2**. Each survey respondent was then asked to assign risk to errors (A, none; B, slight; C, moderate; D, significant; E, dangerous). Risk codes were assigned to each cell in a 5 × 5 grid, with the risk due to correct identification of BG range precoded to "A" (none), as shown in **Table 3**.

The methods and equations for processing the completed forms were as follows: A 10 mg/dl increment master grid (BG from 0–550 mg/dl) was superimposed over each respondent's BG ranges. The average risk score (0–4: 0 = A, 1 = B, … , 4 = E) from the $N = 100$ respondents was computed for each cell in the master grid. A triangular filter was applied to smooth the average scores:

- $Y_{i,j} = 0.25Y_{i-1,j} + 0.50Y_{i,j} + 0.25Y_{i+1,j}$

- $Y_{i,j} = 0.25Y_{i,j-1} + 0.50Y_{i,j} + 0.25Y_{i,j+1}$.

Table 1.
Definition of the Risk Zones and Coordinates of the Lines Defining the Zones of the Parkes Error Grid in a Two-Dimensional Grid[a]

Type 1 diabetes		Type 2 diabetes	
Line of identity			
0/0 550/550		0/0 550/550	
A (clinically accurate measurements, no effect on clinical action)			
B (altered clinical action, little or no effect on clinical outcome)			
B Lower	B Upper	B Lower	B Upper
50/0 50/30 170/145 385/300 550/450	0/50 30/50 140/170 280/380 430/550	50/0 50/30 90/80 330/230 550/450	0/50 30/50 230/330 440/550
C (altered clinical action, likely to affect clinical outcome)			
C Lower	C Upper	C Lower	C Upper
120/0 120/30 260/130 550/250	0/60 30/60 50/80 70/110 260/550	90/0 260/130 550/250	0/60 30/60 280/550
D (altered clinical action, could have significant clinical risk)			
D Lower	D Upper	D Lower	D Upper
250/0 250/40 550/150	0/100 25/100 50/125 80/215 125/550	250/0 250/40 410/110 550/160	0/80 25/80 35/90 125/550
E (altered clinical action, could have dangerous consequences)			
	E Upper		E Upper
	0/150 35/155 50/550		0/200 35/200 50/550

[a] x axis, reference values from 0 to 550 mg/dl; y axis, test device results from 0 to 550 mg/dl.
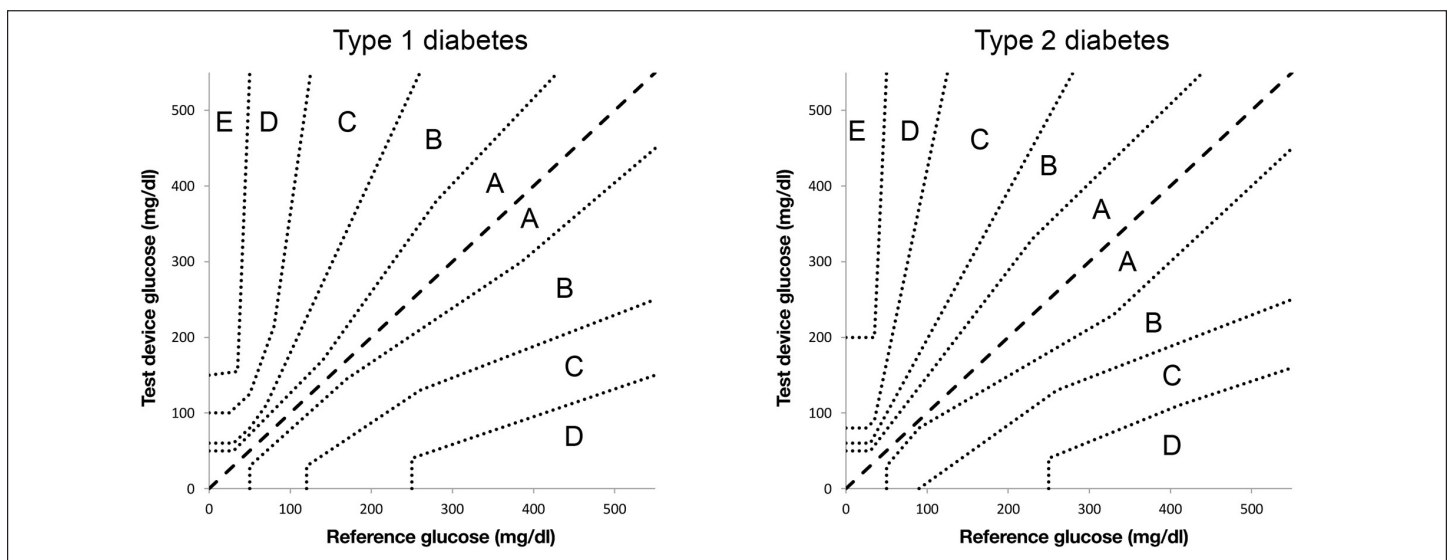

Figure 1. Parkes error grid for patients with **(A)** type 1 diabetes and **(B)** type 2 diabetes. The version for type 1 diabetes is also used for regulatory purposes.

Table 2. Blood Glucose Ranges and Associated Actions	
Code	Action
1	Emergency treatment for *low* BG (e.g., glucagon, intravenous glucose, trip to emergency room)
2	Take glucose (tablets or food)
3	No action needed
4	Take insulin
5	Emergency treatment for *high* BG (visit doctor or hospital; with positive test for ketones)

Resulting error grid zones were based on curves of constant risk. Piecewise linear fits were used to remove uninterpretable oscillations resulting from the filtering in the some of the risk zones.

Zone A is defined in the Parkes error grid as the zone of "clinical accurate measurements with no effect on clinical action."[1] In contrast to the Clarke grid, the Parkes error grid was developed totally independent from ISO15197:2003 criteria (i.e., ±20%[7]) and also specifies a more strict definition for zone B as "altered clinical action with little or no effect on clinical outcome."[1] The Clarke grid specifies zone B as "values that deviate from the reference by >20% but would lead to benign or no treatment based on our assumptions."[2]

Table 3. Degree of Risk Form for Type 1 Diabetes			
Patient's meter reading is in range CODE	Patient is taking the following action	Actual value is in range	Degree of risk (A, B, C, D, or E) *Assign degree of risk below*
1	Emergency treatment for low blood glucose (e.g., glucagon, intravenous glucose, trip to emergency room)	1	A
		2	
		3	
		4	
		5	
2	Take glucose	1	A
		2	
		3	
		4	
		5	
3	No action needed	1	A
		2	
		3	
		4	
		5	
4	Take insulin	1	A
		2	
		3	
		4	
		5	
5	Emergency treatment for high blood glucose (with positive test for ketones)	1	A
		2	
		3	
		4	
		5	

The ISO15197:2003 standard specified that 95% of the values had to be within 20% of the reference value for values ≥75 mg/dl or within ±15mg/dl for values <75 mg/dl,[7] while the new ISO15197 criteria requires 95% of the values to be within 15% of the reference value for values ≥100 mg/dl or within ±15 mg/dl for values <100 mg/dl.[5] It needs to be made clear though that the International Organization for Standardization (ISO) standard defines analytical accuracy and not clinical accuracy.

Many accuracy reports and manuscripts also present data according to Clarke and/or Parkes error grids, usually defining a device with 95% of the data points in zones A or B to meet the grid criteria. This means that 5% of the values can be entirely out of the analytically acceptable range and the BG meter system would still be defined to be an accurate device. However, this is also true of ISO criteria, not just error grids, and has also been adopted for the use of the Parkes grid in the new ISO guideline—currently there is no limit on how far out the other 5% of the values can be to be acceptable.

The performance zone boundaries of the Parkes error grid may appear to be less complicated than the Clarke error grid, because with the Parkes error grid, the zones of inaccuracy fan out from the most accurate in the center to the least accurate at the edges. With the Clarke error grid, there are sharp edges to the boundaries and zone skips. For example, the lower boundaries of the upper D zone are extremely close to the A zone at BG levels between 70 and 110 mg/dl. Even relatively accurate BG meter systems, therefore, may have some data point measurements that do not quite qualify for zone A, and these can become assigned to zone D with no possibility for being assigned to zones B or C. It clearly needs to be emphasized, however, that the Clarke error grid was originally not created to serve as an analytical accuracy standard, but rather to be used as an educational tool for health care professionals and patients. The authors of the Parkes error grid avoided incongruities and skipped zones by defining contiguous boundaries.

Recommendations for Optimal Use of the Parkes Error Grid

In our opinion, the following three procedures should be recommended to all investigators to comply with regulatory and scientific standards:

1. For optimal results, select a reference method that uses the same technology that was used for calibration or that was developed by the same manufacturer, if this is possible. It has been shown that hexokinase-based reference methods and glucose-oxidase-based reference methods can vary by 3–8% from each other.[8,9] However, differences in the underlying measurement technology may also be acceptable, provided that the reference method is calibrated against traceable standards (e.g., National Institute of Standards and Technology).

2. To obtain the paired meter/reference values, perform a clinical study with the protocol and glucose ranges for system accuracy testing as set forth in the ISO15197 guidelines (system accuracy testing, section 7). Unfortunately, it is a common practice to perform accuracy studies that deviate protocol-wise from the ISO guidelines but to analyze the data with the new ISO15197 analysis methods. Studies that perform ISO analysis without following the standard protocol are not allowed to make claims regarding compliance or noncompliance of meters with the ISO15197 standards. We believe that authors of manuscripts intended for medical journals that present such incongruous methods for testing and reporting clinical accuracy of BG accuracy should thoroughly disclose and discuss this weakness in their reporting and analysis. User performance evaluations should be conducted according to ISO15197 (user performance evaluation, section 8) requirements.

3. Use the type 1 diabetes version of the Parkes error grid for system clinical accuracy assessments. The ISO15197 protocols for system accuracy testing are usually performed with blood samples from persons with type 1 diabetes and from persons with type 2 diabetes, as well as with blood samples from subjects who do not have diabetes. The type 1 diabetes version of Parkes error grid has stricter boundaries than the type 2 diabetes version of the Parkes error grid. Because meters may be used by persons with either type of diabetes, the stricter (or less forgiving) type 1 diabetes version of the error grid should therefore be the analysis method of choice for accuracy assessments.

Further Aspects

Many clinical experts in BG meter performance believe that, based on the definition of the risk boundaries, a clinically accurate meter should show at least 95% of its data points in zone A of the Parkes grid (as opposed to 95% of values in either zone A or zone B). It should be noted that, as of the time this article was written, no regulatory body has specified performance standards that include the percentage of data points that must fall within a particular zone or zones of any error grid. Some manufacturers have elected to apply the Parkes metric to their products by specifying that data points in both the A and B zones are clinically acceptable. However, the Parkes original error grid article[1] and our understanding of the opinions of many leading clinicians currently do not specifically support that claim. Furthermore, there appears to be the opinion in the clinical chemistry and regulatory communities that points in the A zone are clinically acceptable but points in the B zone and all the other zones are not clinically acceptable for BG monitors.

The ISO15197:2013 guideline specifies use of the Parkes error grid for assessing outlier data points that do not meet the analytical accuracy requirements. In the published version of this new guideline, 95% of the data points must be analytically acceptable.[7] The Parkes error grid will be used here to possibly salvage some of the analytically unacceptable data points (because they are too widely divergent from the reference measurement) if they are in the A zone of clinical accuracy. Therefore, since ISO15197:2013 is intended to specify analytical accuracy and Parkes is intended to specify clinical accuracy, the inclusion of the Parkes metric in this new ISO guideline is a recognition that some analytically unacceptable BG monitor data points might still be clinically acceptable. It is nevertheless not appropriate in most cases to use this metric, which was developed for quantifying clinical accuracy, to describe analytical accuracy of BG monitor data points.

Modern branded BG meter systems have reached higher analytical accuracy performance levels. Pfützner and coauthors[10] have shown that all flagship meters of the larger manufacturers meet the ISO15197:2003 standards, and 100% of all paired values (i.e., meter/reference) are in zone A of the Parkes error grid.

Currently, a new method for error grid analysis is under development by the Food and Drug Administration, Diabetes Technology Society, the American Diabetes Association, The Endocrine Society, and the Association for Advancement of Medical Instrumentation. This modern error grid will provide guidance about the clinical performance requirements for BG meters. It remains to be seen how much this new error grid will deviate from the Parkes error grid and how it will be applied.

Until such time that a new error grid is agreed upon and published, this article may be useful to many since it provides the necessary information (i.e., coordinates) to allow the general public to create both type 1 and type 2 diabetes Parkes error grids.

References:

1. Parkes JL, Slatin SL, Pardo S, Ginsberg BH. A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose. Diabetes Care. 2000;23(8):1143–8.

2. Clarke WL, Cox D, Gonder-Frederick LA, Carter W, Pohl SL. Evaluating clinical accuracy of systems for self-monitoring of blood glucose. Diabetes Care. 1987;10(5):622–8.

3. Gough DA, Botvinick EL. Reservations on the use of error grid analysis for the validation of blood glucose assays. Diabetes Care. 1997;20(6):1034–6.

4. Cox DJ, Gonder-Frederick LA, Kovatchev BP, Julian DM, Clarke WL. Understanding error grid analysis. Diabetes Care. 1997;20(6):911–2.

5. DIN EN ISO15197: *In vitro* diagnostic test systems -- requirements for blood glucose monitoring systems for self-testing in managing diabetes mellitus (ISO/DIS15197:2010). Draft.

6. Klonoff DC. The need for clinical accuracy guidelines for blood glucose monitors. J Diabetes Sci Technol. 2012;6(1):1–4.

7. ISO15197:2003: *In vitro* diagnostic test systems -- requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus.

8. Twomey PJ. Plasma glucose measurement with the Yellow Springs Glucose 2300 STAT and the Olympus AU640. J Clin Pathol. 2004;57(7):752–4.

9. Genter PM, Ipp E. Accuracy of plasma glucose measurements in the hypoglycemic range. Diabetes Care. 1994;17(6):595–8.

10. Pfützner A, Mitri M, Musholt PB, Sachsenheimer D, Borchert M, Yap A, Forst T. Clinical assessment of the accuracy of blood glucose measurement devices. Curr Med Res Opin. 2012;28(4):525–31.