Published in final edited form as:
Diabetes Care. 2007 February ; 30(2): 449–451. doi:10.2337/dc06-1901.

A Critical Appraisal of the Continuous Glucose—Error Grid Analysis:

Response to Wentholt et al

WILLIAM L. CLARKE, MD[1], **LINDA GONDER-FREDERICK, PHD**[2], **DANIEL COX, PHD**[2], and **BORIS KOVATCHEV, PHD**[2]

1*Department of Pediatrics, University of Virginia Health System, Charlottesville, Virginia*

2*Department of Psychiatry and Neurobehavioral Sciences, University of Virginia Health System, Charlottesville, Virginia.*

In a recent publication, Wentholt et al. (1) stated that their aim was to critically explore the continuous glucose—error grid analysis (CG-EGA) (2) and to compare it with traditional techniques using data previously reported from two sensors. As developers of the CG-EGA, we hoped that our method might stimulate a discussion on the important problem of the accuracy of continuous monitoring sensors (CGS); therefore, we read this critique with interest.

The methods used by Wentholt et al. (1) unfortunately failed to take into account the basic structure of CGS data, which represent time series (i.e., sequential readings that are ordered in time) (3). This structure leads to two fundamental requirements in their analysis. First, consecutive sensor readings taken from the same subject within a relatively short time are highly interdependent. Therefore, standard statistical analyses such as *t* tests, while appropriate for independent data points, will produce inaccurate results if applied to CGS data. Second, the order of the CGS data points is essential for clinical decision making. For example, the sequences 90 → 82 → 72 mg/dl and 72 → 82 → 90 mg/dl are clinically very different. Standard accuracy measures, such as the mean absolute deviation (MAD) used by Wentholt et al. (1), do not account for the data's temporal order; if reference-sensor data pairs are reshuffled, the MAD remains the same.

As a result, the primary statistical analysis used by Wentholt et al. is flawed, both to demonstrate significant differences between the sensors and to imply that CG-EGA is insensitive. The CGS data from 13 subjects were pooled to compare 2 MADs (15.0 ± 12.2 vs. 13.6 ± 10.2%). The result was reported as significant ($P = 0.013$), but for these highly overlapping MADs to differ statistically required a large number (>1,000) of degrees of freedom, which was calculated by pooling the total number of CGS data points (735 and 1,156) across all subjects. Such an approach led to inaccurate conclusions because there were only 13 independent subjects, and the data points within each subject were highly dependent. If the correct number of degrees of freedom is used, the MADs of the two sensors are not different ($P > 0.5$), which confirms the CG-EGA results showing no differences.

Other conclusions by Wentholt et al. also deserve comment. First, they stated that CG-EGA is time consuming. Indeed, analyses of temporal data are intrinsically more sophisticated than standard time-independent statistics, but such analyses are essential for this type of data. CG-

Address correspondence to Boris Kovatchev, Department of Psychiatry and Neurobehavioral Sciences, University of Virginia Health System, Box 800137, Charlottesville, VA 22908. E-mail: E-mail: boris@virginia.edu..

EGA software is available. Second, Wentholt et al. stated that "poor accuracy rate is barely noticeable in the final CG-EGA outcome," implying that this result of the CG-EGA is incorrect. However, this result is not incorrect because better combined (rate and point) accuracy during hypoglycemia is observed with the sensor, showing poorer rate accuracy in this critical region. It is clinically apparent that when blood glucose is <3.9 mmol/l point accuracy should be given more emphasis than rate accuracy. A strength of CG-EGA is its ability to vary the input of either rate or point accuracy to overall clinical accuracy depending on blood glucose range. Third, the results of CG-EGA vary with time intervals. This is also an intuitive strength of CG-EGA, which is designed to account for increased noise associated with frequent sampling. We advocated (2) adopting a uniform sampling protocol with reference and/or sensor pairs taken every 10–15 min to standardize comparisons of rate accuracy, which is a sampling scheme based on physiological considerations of possible glucose change rates. Fourth, Wentholt et al. (1) questioned the appropriateness of the formulae to shift point EGA based on interstitial time lag. However, the authors reported an average time lag of ~7 min in one of their sensors, which is identical to that assumed for CG-EGA, thus confirming that ~7 min is a reasonable average for blood-to-interstitial diffusion delays. CG-EGA software allows setting this parameter to any value <7 min.

We are pleased that both the discussion regarding CG-EGA and the analysis of time series data have begun, and we look forward to continuing this important dialogue. However, we also recommend careful consideration of basic statistical assumptions when analyzing sensor-generated glucose data; their inherent temporal structure should be taken into account.

References

1. Wentholt IM, Hoekstra JB, DeVries JH. A critical appraisal of the continuous glucose—error grid analysis. Diabetes Care 2006;29:1805–1811. [PubMed: 16873784]
2. Kovatchev BP, Gonder-Frederick LA, Cox DJ, Clarke WL. Evaluating the accuracy of continuous glucose-monitoring sensors: continuous glucose—error grid analysis illustrated by Therasense Freestyle Navigator data. Diabetes Care 2004;27:1922–1928. [PubMed: 15277418]
3. Chatfield, C. The Analysis of Time Series: An Introduction. Vol. 6th ed. Chapman and Hall/CRC; Boca Raton, FL: 2004. p. 1