Dynamical glucometry: Use of multiscale entropy analysis in diabetes

Madalena D. Costa,^{1,2} Teresa Henriques,^{1,3} Medha N. Munshi,^{2,4} Alissa R. Segal,^{4,5} and Ary L. Goldberger^{1,2}

¹*The Wyss Institute for Biologically Inspired Engineering at Harvard University, Boston, Massachusetts 02115, USA*

²*Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215, USA*

³*Center for Research in Health Technologies and Information Systems (CINTESIS), Porto, Portugal and Instituto de Telecomunicações, Porto, Portugal*

⁴*Joslin Diabetes Center, Boston, Massachusetts 02215, USA*

⁵*MCPHS University, Boston, Massachusetts 02215, USA*

(Received 25 April 2014; accepted 22 August 2014; published online 23 September 2014)

Diabetes mellitus (DM) is one of the world's most prevalent medical conditions. Contemporary management focuses on lowering mean blood glucose values toward a normal range, but largely ignores the dynamics of glucose fluctuations. We probed analyte time series obtained from continuous glucose monitor (CGM) sensors. We show that the fluctuations in CGM values sampled every 5 min are not uncorrelated noise. Next, using multiscale entropy analysis, we quantified the complexity of the temporal structure of the CGM time series from a group of elderly subjects with type 2 DM and age-matched controls. We further probed the structure of these CGM time series using detrended fluctuation analysis. Our findings indicate that the dynamics of glucose fluctuations from control subjects are more complex than those of subjects with type 2 DM over time scales ranging from about 5 min to 5 h. These findings support consideration of a new framework, *dynamical glucometry*, to guide mechanistic research and to help assess and compare therapeutic interventions, which should enhance complexity of glucose fluctuations and not just lower mean and variance of blood glucose levels. © 2014 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4894537>]

Patients with type II diabetes mellitus (DM) need to closely monitor their blood glucose levels to manage the disease and help avoid its associated problems. Continuous glucose monitoring (CGM) systems were introduced into clinical practice approximately a decade ago. Using a sensor inserted under the skin, these systems provide glucose levels in tissue fluid at 5 min intervals for up to a week. The information extracted from these recordings for clinical use is primarily based on measures of the mean and range of glucose variability. One opportunity to improve our understanding and therapy of this highly prevalent disease is by looking for information encoded in the dynamical structure of glucose fluctuations. In this study, using techniques from statistical physics and nonlinear dynamics, we probed the multiscale complexity of glucose fluctuations obtained with CGM devices from elderly subjects with type II DM and from non-diabetic control subjects. We found that the glucose fluctuations in the patients with DM were significantly less complex compared to those in control subjects. Furthermore, we found that this loss of complexity could not be accounted for simply by changes in the average or range of glucose values. These findings are of interest because they raise the possibility of using information in glucose time series (“dynamical glucometry”) in both laboratory and clinical settings. In particular, these findings support the concept that diagnostic and therapeutic assessments should not only take into account the mean and variance of glucose values but also measurement of properties related to their fluctuations over multiple time scales.

INTRODUCTION

Blood glucose control, central to physiologic fuel regulation, involves the complex interplay of hormones (e.g., insulin and glucagon), food intake, muscle activity, liver, and fat cell function, among other factors.¹ Type 2 DM is marked by hyperglycemia due to tissue resistance to insulin and to relative insufficiency of insulin secretion. The disease affects tens of millions worldwide. Complications include stroke, heart attack, blindness, limb loss, neurologic dysfunction, and kidney failure. Because of its association with obesity, the incidence of type 2 DM is increasing at an alarming rate. Current management relies on diet and exercise as well as drugs that lower blood sugar toward a normal range, including insulin therapy.² However, type 2 DM remains a major challenge in terms of prevention and effective treatment.

The most widely used way of guiding day-to-day therapy is with blood glucose measurements obtained by finger sticks. In addition, the hemoglobin A1c (HbA1c) blood test provides information about glucose levels over the preceding several months.^{1,2} A more recent and minimally invasive modality uses CGM technology, with sensors implanted subcutaneously for about a week.² However, the time series derived from CGM recordings remain a largely untapped source of dynamical information. In clinical practice, the CGM indices of interest are primarily based on calculations of the percentage of time above and below given thresholds, range, and average values of blood glucose. In a small number of research studies,^{3–6} which investigated glucose dynamics using detrended fluctuation analysis (DFA), no

differences were noted between diabetic and non-diabetic subjects on relatively short (5–60 min) time scales.

This communication advances an alternative statistical physics approach to the analysis of CGM time series, primarily based on a complexity measure termed as multiscale entropy (MSE).⁷ The overall goal of this study is to extract information about the nature of physiologic vs. pathologic glucose control from CGM recordings. The specific hypotheses are that: (1) the noisy-appearing fluctuations in blood glucose levels in non-diabetic control subjects are complex, i.e., they are distinct from uncorrelated randomness on scales ranging for minutes to hours, and (2) blood glucose fluctuations are less complex in patients with type 2 DM compared with non-diabetic subjects.

METHODS

We retrospectively analyzed CGM data from 18 patients with type 2 DM and 12 elderly control subjects without DM, enrolled in studies (approved by the Institutional Review Board) conducted at the Joslin Diabetes Center by the Geriatric Diabetes Research Group (Table I). The CGM data were obtained using the iProTM system version 1 or 2 (Medtronic, Inc., Minneapolis, MN), set at a sampling rate of one measurement every 5 min. The duration of the recordings varied between two and four days, therefore, the time series comprised between approximately 550 and 1100 CGM values. All patients with diabetes were on an insulin regimen with/without additional oral anti-diabetic agents. Fig. 1 shows selective CGM time series from a control subject and a patient with type 2 DM.

Results, unless otherwise stated, are reported as means $\pm$ standard deviation (SD). Group values were compared using a two-tailed, non-parametric (Mann–Whitney) test. Statistical significance was set at a p-value $< +0.05$. Coefficients of determination, r^2 , derived from the linear regression of the dynamical variables (complexity index and α -exponent) against standard clinical measures of glucose control (HbA1C, means, and SD of CGM time series) were computed.

RESULTS

To assess whether the short-term (5-min) fluctuations in CGM time series were solely attributable to uncorrelated noise (instrumental and other), we first compared the structure of the time series of the differences between consecutive glucose values, termed first difference time series, with that of their shuffled surrogate time series⁸ (Fig. 2). If the former

TABLE I. Characteristics of control subjects and patients with type 2 DM. Data are expressed as mean $\pm$ SD.

Variables	Non-diabetic	Type 2 DM
n	12	18
Gender (M/F)	3/9	6/12
HbA1c (%)	5.6 $\pm$ 0.2	6.8 $\pm$ 0.8
Age (yr)	79.3 $\pm$ 5.0	76.0 $\pm$ 5.0
BMI (lb/in ²)	25.8 $\pm$ 6.93	31.0 $\pm$ 7.0

FIG. 1. Time series of blood glucose derived from the CGM recordings of a non-diabetic control subject (a) and a patient with type 2 diabetes (b). The inset is a magnification of a 24-h data segment.

were less entropic than the latter, one would infer that the structure of the original first difference time series was distinct from that of uncorrelated randomness and, consequently, that there was information in the short-term CGM fluctuations.

Thus, for each first difference time series, we generated 100 surrogate time series, obtained by shuffling the order of the data points in the original series. To quantify entropy, i.e., the degree of irregularity of the time series, we used the sample entropy (SampEn) algorithm.⁹ SampEn is a conditional probability measure quantifying the likelihood that if a sequence of m consecutive data points matches (within a tolerance r) another sequence of the same length, then the two will still match when their length increases from m to $m + 1$ data points. Here, SampEn was calculated with standard parameter values⁹ $m = 2$ and $r = 0.15$. Regular and/or periodic time series have (theoretically) zero entropy and uncorrelated random time series have maximum entropy, whose specific value depends on the time series length.⁹

Of note, using a percentage of a time series' SD as the r value for the computation of SampEn is equivalent to normalizing the time series to unit variance and zero mean prior to the computation of SampEn. Therefore, differences in SampEn are not attributable to differences in the mean or SD of the signals.

FIG. 2. Time series of the differences between consecutive blood glucose values derived from the CGM recording of a control subject (a) and a surrogate signal (b) obtained by shuffling the temporal order in which these difference values appear.

Figure 3 shows that for each control subject, the 100 shuffled surrogate time series were consistently more entropic than the original first difference time series. Qualitatively similar results were obtained for the first difference time series of patients with type 2 DM.

These findings support the hypothesis that the short-term fluctuations in CGM values do not represent uncorrelated randomness. Accordingly, we next investigated the question of whether the glucose fluctuations over a range of time scales were more complex in control vs. diabetic subjects. For this purpose, we employed the MSE method,^{7,10–12} which quantifies the degree of irregularity of a time series over a range of time scales. Briefly, the method comprises three steps: (1) using the original signal to derive a set of time series representing the system's dynamics on different time scales through a coarse-graining procedure, (2) calculating sample entropy (SampEn)⁹ for each coarse-grained time series, and (3) computing a complexity index, defined here as the sum of SampEn values over a pre-defined range of scales. The coarse-grained time series for scale i is obtained by averaging the data points inside consecutive non-overlapping windows of length i . The MSE curves are obtained by plotting the SampEn values for the coarse-grained time series as a function of scale. In this study, given the length of the recordings and the fact that glucose values were recorded every 5 min, we explored scales 1 to 6, where scale 1 corresponds to 5 min and scale 6 to 30 (5×6) min. With $m=2$ ($m+1=3$), the SampEn values of the coarse-grained time series for scale 6, probe glucose fluctuation patterns that range from 60 (2×30) to 90 (3×30) min.

The complexity index derived from the MSE analysis of the glucose dynamics on scales 1 to 6 (Fig. 4) was significantly higher ($p < 0.004$) for the non-diabetic than the diabetic group (Fig. 5, left panel). On time scales ranging from 1 to 12, the complexity index remained significantly higher ($p < 0.007$) for the non-diabetic than the diabetic group. Since the number of data points in the coarse-grained time series for scale 12 is only half of that in the coarse-grained time series for scale 6, we used a less restrictive r value (0.25) in this analysis to avoid spuriously high entropy values due to finding none or very few matches. The results obtained are independent of differences between the two groups in the mean and SD values of the CGM time series.

To further characterize the dynamics of CGM time series and, in particular, to investigate their long-range correlation properties, we computed the fractal scaling exponents

FIG. 3. SampEn values for the first difference time series (solid circles) of non-diabetic control subjects and their shuffled surrogate (open circles) time series ($n = 100/\text{subject}$).

FIG. 4. MSE analysis of glucose time series for the non-diabetic ($n = 12$) and diabetic ($n = 18$) groups. Symbols represent the mean values of entropy for each group at each scale and the bars represent the SD.

using the DFA method.¹³ Briefly, the algorithm, a modified root mean square analysis of a random walk, quantifies the relationship between $F(n)$, the root mean square fluctuation of an integrated and detrended time series, and the observation window size, n . $F(n)$ increases with window size according to $F(n) \propto n^\alpha$. If $\alpha = 0.5$, there are no correlations and the signal is white noise, i.e., uncorrelated randomness. If $0.5 < \alpha < 1$, the signal exhibit long-range temporal correlations. For the special case when $\alpha \simeq 1$ ($1/f$ noise), the time series exhibits scale-invariant properties. If $\alpha = 1.5$, the time series represents a random walk (brown noise). α values ≥ 1.5 are associated with strong trends (“black noise”).

For time scales shorter than 1 h, there were no significant difference between the α -exponents of CGM time series from diabetic (1.87 ± 0.29) and non-diabetic subjects (1.86 ± 0.22). These results are consistent with those previously reported^{3,4} and indicate that the structure of the CGM time series on short time scales is very smooth and characterized, primarily, by monotonically increasing or decreasing trends. In contrast, we found that for time scales larger than 1 h, (Fig. 5, right panel), the α -exponents were significantly ($p < 0.0001$) lower (1.19 ± 0.13), and closer to 1, in control subjects than in patients with type 2 DM (1.42 ± 0.09), respectively. These results indicate that the dynamics of glucose control in non-diabetic subjects over the specified range is fractal, i.e., time-invariant, whereas in this group of subjects with type 2 DM it is not different from brown noise.^{3,4}

FIG. 5. Complexity index (sum of SampEn values for scales 1 to 6, inclusively) and DFA α -exponent (time scales > 1 h) for the non-diabetic and diabetic groups. Symbols with error bars represent mean and SD values, respectively.

TABLE II. Slope m , coefficient of determination, r^2 , and the two-tail p value¹⁴ derived from the linear regression of the dynamical variables, complexity index (CI, scales 1 to 6) and DFA α -exponent for time scales >1 h, against three clinical measures, HbA1c, mean, and SD values of the CGM time series. (See Fig. 6 for graphs.)

Variables	m	r^2	p value
CI vs HbA1c	-0.796	0.19	1.5×10^{-2}
CI vs mean	-0.030	0.33	8.3×10^{-4}
CI vs SD	-0.054	0.34	6.6×10^{-4}
α vs HbA1c	0.110	0.39	2.3×10^{-4}
α vs mean	0.003	0.41	1.3×10^{-4}
α vs SD	0.006	0.52	$<10^{-5}$

Next, we investigated the relationship between the two dynamical variables of interest, the complexity index and the DFA α -exponent, and three clinical (static) measures of glucose control, HbA1c, mean, and SD values of CGM time series (Table II, Fig. 6). We found significant correlations between higher glucose complexity indexes and lower (i.e., better) values of standard clinical measures of diabetes control (lower HbA1c and lower mean and SD of CGM glucose levels). We also found significant correlations between higher DFA α -exponents for time scales >1 h (i.e., those closer to 1.5 or brown noise) and poorer indexes of clinical glucose control. Alternatively stated, the DFA α -values closer to 1 were associated with better clinical control.

DISCUSSION

Our study is notable for a number of reasons. We find that the short-term fluctuations in serum glucose are not white noise (uncorrelated randomness), as shown by the analysis summarized in Fig. 3. Previous attempts to analyze the dynamical structure of the CGM time series on time scales <1 h using DFA failed to discriminate between diabetic and non-diabetic groups. This apparent lack of sensitivity may be due

to the confounding influence of certain types of nonstationarities on DFA computations.¹⁵

We report, using MSE analysis for the first time to our knowledge, that the complexity of the CGM time series is significantly higher for the non-diabetic than the diabetic subjects. From a mathematical point of view, these results indicate that the degree of irregularity of the CGM time series is higher for the non-diabetic than the diabetic group, across a physiologically relevant range of time scales (5–90 min). The reduction in complexity with diabetes cannot be attributed to higher CGM values or higher SD of the CGM time signals. From a biologic point of view, these results, albeit part of a small, retrospective analysis, are consistent with the complexity-loss framework of disease and aging,^{10–12,16} which subtends the notion that a breakdown in underlying physiologic control mechanisms results in less complex, i.e., either more random or more periodic, output signals.

Assessment of the nature of the correlations in the CGM time series, using DFA, on time scales larger or equal to 1 h, confirms and extends previous findings³ of a significant increase in the α -exponent towards that of brown noise (1.5) with type 2 DM. These results are consistent with a loss of fractality in the dynamics of CGM signals. Furthermore, our findings are based on 2–4 day recordings, not solely on a single day of CGM data as previously reported.³

Dynamical and conventional assays of glucose control are complementary. Higher complexity indices and fractal exponents (close to 1) were correlated with more salutary clinical measures of both short and long term glucose control, i.e., lower glucose mean and SD values and lower HbA1c. However, the correlations were modest, consistent with the fact that dynamical and the widely used clinical measures are not redundant.

Current therapeutic strategies focus almost exclusively on glucose lowering by searching for molecular “targets”

FIG. 6. Linear regression of dynamical (complexity index (scales 1 to 6) and DFA α -exponent for time scales >1 h) against clinical (HbA1c, mean, and SD values of the CGM time series) measures of glucose control. Statistical summaries are presented in Table II.

involved in increasing sensitivity to insulin or insulin release by the pancreas. Our findings support extending the conventional notion of a discrete (pathway or signaling) therapeutic target to include the overall control “system as target.” For example, two interventions may reduce blood glucose to a similar extent; however, one may do so by restoring healthy multiscale dynamics, the other may fail to do so. We anticipate that the former would be preferable to the latter. Furthermore, we suggest that dynamical glucometry, defined here as the application of an ensemble of computational tools, including multiscale entropy, to the analysis of serum glucose time series will help uncover hidden information of basic and clinical value.

We note that CGM time series obtained with greater precision and with higher sampling frequency than those available here may provide more information about glucose control in health and disease. The sampling rate employed in current clinical CGM studies obviates the possibility of probing glucose dynamic on scales shorter than 5 min, which earlier studies suggested are physiologically important.^{17,18}

Future prospective studies may also help address additional questions, including: are serum glucose fluctuations in younger healthy subject more complex than in healthy elderly? Is loss of complexity associated with increased comorbidity related to multiorgan dysfunction? The concomitant use of continuous heart rate monitoring in assessing diabetic therapeutics might add additional information about neuro-autonomic function. Multimodal serial measurements of insulin and other molecular regulators involved in physiologic fuel control could also contribute important information.

Finally, our findings, which confirm and extend previous work,^{3–6,19–22} may be useful in developing and testing mathematical models of physiologic fuel regulation in health and disease.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health Grants Nos. K99/R00 AG030677 (M.D.C.) and R01GM104987 (A.L.G.); the G. Harold and Leila Y. Mathers Charitable Foundation (A.L.G.), the James S.

McDonnell Foundation (M.D.C.); the Portuguese National Science Foundation through the Ph.D. Grant No. SFRH/BD/70858/2010 and the Graetz Foundation (M.M.).

- ¹American Diabetes Association, *Diabetes Care* **37**(Suppl. 1), S81 (2014).
- ²American Diabetes Association, *Diabetes Care* **37**(Suppl. 1), S14 (2014).
- ³H. Ogata, K. Tokuyama, S. Nagasaka, A. Ando, I. Kusaka, N. Sato, A. Goto, S. Ishibashi, K. Kiyono, Z. R. Struzik *et al.*, *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **291**, R1638 (2006).
- ⁴H. Ogata, K. Tokuyama, S. Nagasaka, A. Ando, I. Kusaka, N. Sato, A. Goto, S. Ishibashi, K. Kiyono, Z. R. Struzik *et al.*, *Methods Inform. Med.* **46**, 222 (2007).
- ⁵N. Yamamoto, Y. Kubo, K. Ishizawa, G. Kim, T. Moriya, T. Yamanouchi, and K. Otsuka, *Diabetes Technol. Ther.* **12**, 775 (2010).
- ⁶H. Ogata, K. Tokuyama, S. Nagasaka, T. Tsuchita, I. Kusaka, S. Ishibashi, H. Suzuki, N. Yamada, K. Hamano, K. Kiyono *et al.*, *Metabolism* **61**, 1041 (2012).
- ⁷M. Costa, A. L. Goldberger, and C.-K. Peng, *Phys. Rev. Lett.* **89**, 068102 (2002).
- ⁸J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, *Physica D* **58**, 77 (1992).
- ⁹J. S. Richman and J. R. Moorman, *Am. J. Physiol. Heart Circ. Physiol.* **278**, H2039 (2000).
- ¹⁰M. Costa, I. Ghiran, C.-K. Peng, A. Nicholson-Weller, and A. L. Goldberger, *Phys. Rev. E* **78**, 20901 (2008).
- ¹¹L. C. Cancio, A. I. Batchinsky, W. L. Baker, C. Necsoiu, J. Salinas, A. L. Goldberger, and M. D. Costa, *J. Crit. Care* **28**, 1093 (2013).
- ¹²M. D. Costa, W. T. Schnettler, C. Amorim-Costa, J. Bernardes, A. Costa, A. L. Goldberger, and D. A. de Campos, *Early Hum. Dev.* **90**, 67 (2014).
- ¹³C.-K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger, *Chaos* **5**, 82 (1995).
- ¹⁴P. Wessa, Linear regression graphical model validation (v1.0.7) in free statistics software (v1.1.23-r7). Office for Research Development and Education (2012), see http://www.wessa.net/rwasp_linear_regression.wasp/.
- ¹⁵K. Hu, P. C. Ivanov, Z. Chen, P. Carpena, and H. E. Stanley, *Phys. Rev. E* **64**, 011114 (2001).
- ¹⁶M. Costa, A. L. Goldberger, and C.-K. Peng, *Phys. Rev. E* **71**, 021906 (2005).
- ¹⁷G. E. Anderson, Y. Kologlu, and C. Papadopoulos, *Metabolism* **16**, 586 (1967).
- ¹⁸A. Iberall, M. Ehrenberg, S. Cardon, and M. Simenhoff, *Metabolism* **17**, 1119 (1968).
- ¹⁹J. Churrua, L. Vigil, E. Luna, J. Ruiz-Galiana, and M. Varela, *Diabetes Metab. Syndr. Obes.* **1**, 3 (2008).
- ²⁰R. Brunner, G. Adelsmayr, H. Herkner, C. Madl, and U. Holzinger, *Crit Care* **16**, R175 (2012).
- ²¹M. Signal, F. Thomas, G. M. Shaw, and J. G. Chase, *J. Diabetes Sci. Technol.* **7**, 1492 (2013).
- ²²R. T. M. van Hooijdonk, A. Abu-Hanna, and M. J. Schultz, *Crit Care* **16**, 178 (2012).