DTT

Diabetes Technology & Therapeutics

ORIGINAL ARTICLE

Mobile Phone-Based Pattern Recognition and Data Analysis for Patients with Type 1 Diabetes

Stein Olav Skrøvseth, Ph.D.,[1] Eirik Årsand, Ph.D.,[1,2] Fred Godtliebsen, Ph.D.,[3]
and Gunnar Hartvigsen, Ph.D[1,2]

Abstract

Background: Persons with type 1 diabetes who use electronic self-help tools, most commonly blood glucose meters, record a large amount of data about their personal condition. Mobile phones are powerful and ubiquitous computers that have a potential for data analysis, and the purpose of this study is to explore how self-gathered data can help users improve their blood glucose management.
Subjects and Methods: Thirty patients with insulin-regulated type 1 diabetes were equipped with a mobile phone application for 3–6 months, recording blood glucose, insulin, dietary information, physical activity, and disease symptoms. The data were analyzed in terms of usage of the different modules and which data processing and visualization tools could be constructed to support the use of these data.
Results: Eighteen patients (denoted "adopters") recorded complete data for over 80 consecutive days, up to 247 days. Among those who withdrew or did not use the application extensively, the most common reasons given were outdated or difficult-to-use phone. Data analysis using period finding and scale-space trends was found to yield significant patterns for most adopters. Pattern recognition methods to predict low or high blood glucose were found to be performing poorly.
Conclusions: Minimally intrusive mobile applications enable users with type 1 diabetes to record data that can provide data-driven feedback to the user, potentially providing relevant insight into their disease.

Introduction

$\mathbf{T}$YPE 1 DIABETES (T1D) is a chronic disease that requires considerable effort by the afflicted patients to control their blood glucose level. Recent technological advances have provided better devices and tools, bringing patients closer to closed-loop and artificial pancreas solutions.[1,2] Nevertheless, a large percentage of patients still use paper-based methods to track their blood glucose and related factors in multiple daily injections of insulin regimens combined with self-measured blood glucose (SMBG).

Physiological and metabolic models of the interactions among blood glucose, insulin, diet, and other factors are useful as decision tools for patients.[3] Computational methods including mathematical models tend to be based on continuous glucose meters (CGMs), which measure the glucose level in the subcutaneous tissue. Because CGMs do not measure in the blood directly, there is a significant delay and more inaccurate measurements, particularly during critical periods of quickly decreasing blood glucose levels or with improper calibration or sensor fixation. The vast amount of data provided by CGMs provides rich datasets for modeling. Pro-

viding patients with CGMs seems to enable better glycemic control and lower glycated hemoglobin (HbA1c), particularly in patients with high baseline HbA1c.[3,4] Compared with the simplicity of SMBG, however, CGMs are a very resource-intensive and expensive solution for many patients.

Glycemic variability is considered an important clinical variable for T1D patients.[5] There are several measures of variability, such as the commonly used mean amplitude of glycemic excursions. Quantifying variability using SMBG values is typically done using SD even though the values are generally not normally distributed. The correspondences between the different measures of variability are not clear.[6]

Using information technologies as tools for patients with chronic diseases is effective if they are used properly; in particular, intelligent decision support systems have proved to be promising.[7] Numerous tools using mobile technology to support patients with diabetes exist, either alone or in combination with online services.[8,9] Using mobile phones as a communications tool for people with diabetes has been shown effective,[10] and using mobile phone technology in conjunction with telemedical support improves HbA1c levels.[11] Given the large amount of data assembled by T1D

[1]Norwegian Centre for Integrated Care and Telemedicine, University Hospital of North Norway, Tromsø, Norway.
Departments of [2]Computer Science and [3]Mathematics and Statistics, University of Tromsø, Tromsø, Norway.
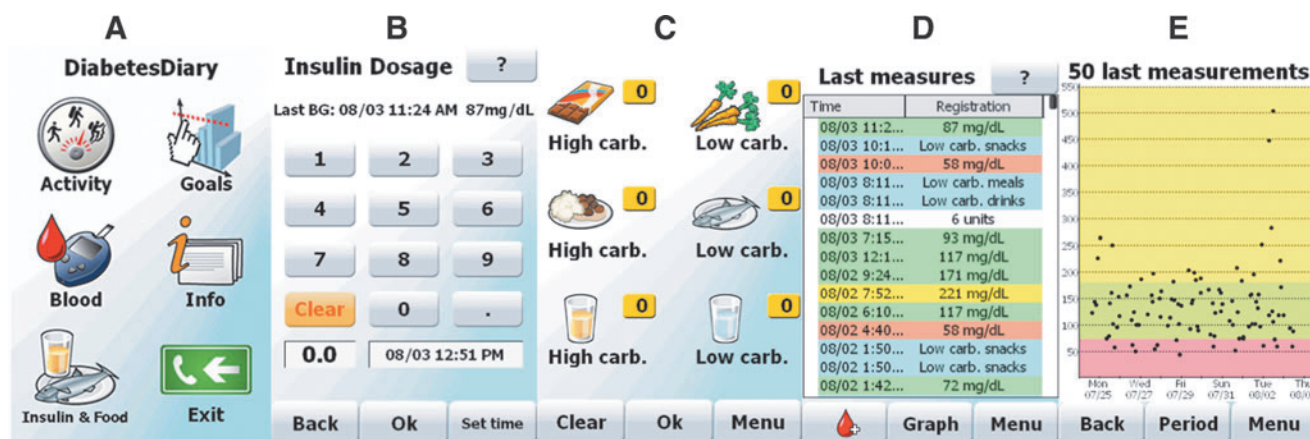
FIG. 1. Screenshots from the Few Touch Application. From left to right, **(A)** the main screen with all basic options, **(B)** recording of insulin dose with editable time (default current time), **(C)** Registration of food and drink into six categories, **(D)** overview list of last activities, with normal blood glucose in green, high glucose in yellow, low glucose in red, food or drink in blue, and insulin in white, and **(E)** graph with the last glucose measurements.

patients and the ubiquitous nature of these devices, utilizing phones' computational power for data analysis and visualization is a natural step forward.[9]

At the Norwegian Centre for Integrated Care and Telemedicine (Tromsø, Norway), a diabetes diary known as the Few Touch Application (FTA) has been developed as a research tool for self-management intended for both T1D and type 2 diabetes (T2D) patients. The version for T2D is mature and has been tested both in pilot groups and in a large-scale randomized controlled trial.[12] The version for T1D used in the present study includes recording of insulin, symptoms, and comments but is otherwise identical to the T2D version.[13] In this open-ended study, patients were encouraged to use the diary as part of their daily life, thus accumulating a realistic dataset. The data from these patients were assembled and analyzed with the aim of identifying which data analysis methods could realistically be used and could provide meaningful feedback to the user.

Determining relevant clinical patterns for patients can provide them with meaningful insight into relevant parameters of their own disease and provide a foundation for discussion with healthcare personnel. Predictive models based on CGMs are widespread, whereas similar models for SMBG are rare because of the relative sparsity of the data. Detecting patterns that lead to unwanted situations can provide useful feedback to patients so that these situations can be avoided.

Trend detection is another potentially useful modality of data-driven feedback. A user with a mean blood glucose level that is too high should ideally change this situation by achieving a long-term decreasing glucose trend. Typically, users lack tools to detect such patterns and to associate them with personal lifestyle or habits.

Subjects and Methods

We developed a version of the FTA for T1D based on previous experience with T2D versions and prototyping and from a previous T1D study.[14] The T1D version comprises a function for easy recording of insulin injections together with nutrition, options for commenting on all data, and an improved feedback screen where the user sees blood glucose,

insulin injections, and food intake in relation to each other. The application was developed on the Windows mobile version 6.5 operating system, and the patients were given an HTC (Taoyuan City, Taiwan) HD mini mobile phone for use during the trial. In Figure 1 we show some screenshots from the application used. The system for recording food intake consisted of six categories of food: low carbohydrate snack, high carbohydrate snack, low carbohydrate drink, high carbohydrate drink, low carbohydrate meal, and high carbohydrate meal (see Fig. 1C for the user interface).

For the study we recruited 30 persons with T1D who were attending the Department of Endocrinology at the University Hospital of North Norway (Tromsø). The recruitment was performed at the discretion of the diabetes nurses and medical doctors. Five of the recruited patients used CGMs along with the FTA.

The patients were instructed to use the application in the way that suited their needs and preferences but were encouraged to measure blood glucose at least three times a day, to record their insulin injections, their nutritional intake, their physical activity, and symptoms of other diseases or discomfort, and to set personal goals for nutrition and physical activity. The patients were invited to meetings at 1, 3, and 6 months (unless they withdrew from the study), where we downloaded data from their mobile phone diaries and the patients gave their feedback on the user interface and other aspects of the application, blood glucose measuring system, and the mobile phone.

Analysis

The data were analyzed in terms of usage using regression coefficients to test for significant characteristics correlating with patient's usage of the FTA.

The recorded data were mined to find useful methods for modeling and prediction. The identified methods and their background are reported in Results.

All statistical analysis was done using R.[15]

Results

Of the 30 patients recruited, three actively withdrew from the study before 3 months. Two of these individuals provided
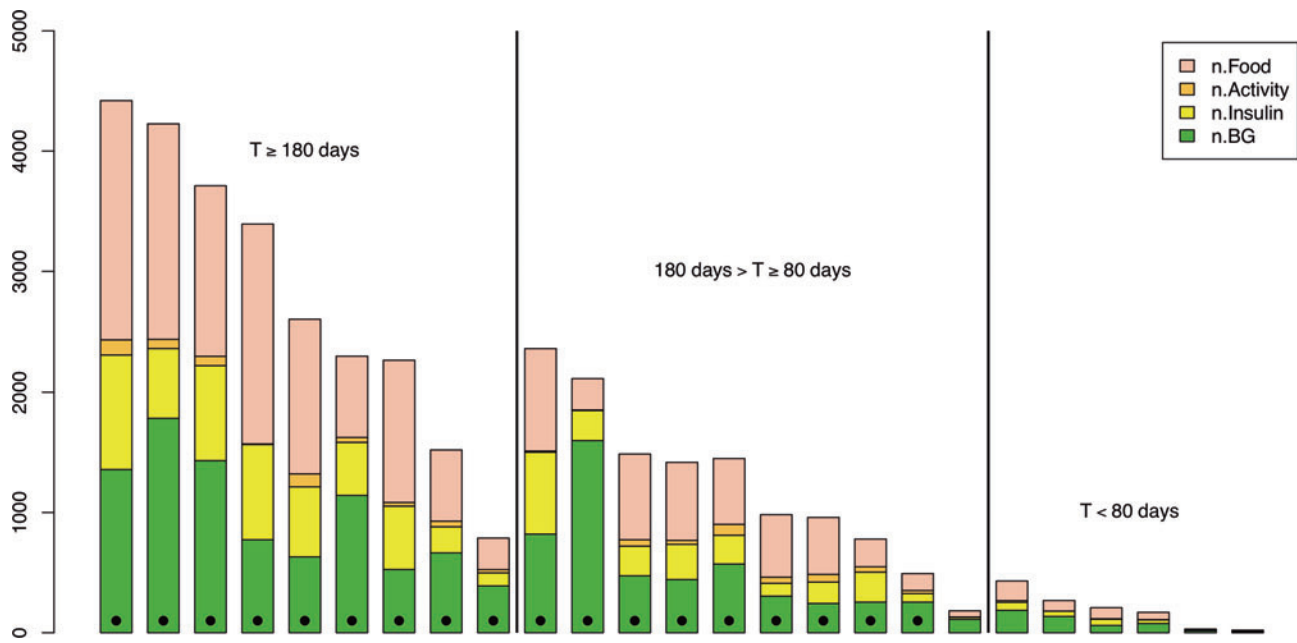
FIG. 2. Usage statistics for all recruited patients who contributed any data. Each column is, from the bottom to the top, the number of recordings in the logs for blood glucose (n.BG), insulin, activity, and food, respectively. The patients classified as adopters are denoted with a dot at the low end of the column. The patients are grouped into three groups according to overall period of usage time *T* as measured from first to last entry in the Few Touch Application. Five users with no data are not shown. Color graphics available online at www.liebertonline.com/dia

the data they had recorded up to the time of withdrawal. Three additional patients' data were lost to the study for various reasons, meaning that 24 patients followed through to the 3-month follow-up. All patients were given the opportunity to continue using the equipment for a further 3-month period; 19 patients accepted this and were followed up for a total of 6 months. Not all patients used the equipment actively during the period they had it.

Usage

As a measure of overall usage, we used the total number of blood glucose measurements, insulin injections recorded, and food items recorded in the period. The results for each patient are summarized in Figure 2. Because physical activity and symptoms were recorded to a significantly lower degree (compare Fig. 2), we did not use these data in the analysis. To test for covariates we used a linear regression model on sex, age, and prestudy HbA1c. Results show that usage was positively correlated with age (regression coefficient, 55.8 recordings/year; $P = 0.009$). Usage was not significantly correlated with prestudy HbA1c ($P = 0.34$) or sex ($P = 0.10$).

There was no change in mean HbA1c for the cohort or any correlation between HbA1c and any parameters of usage. There was no significant change in weight over the course of the study or any correlation between change in weight and usage parameters.

For further analysis, we divided the patients into two disjoint groups: "adopters" and "nonadopters." The adopters were 18 patients who recorded data reliably (i.e., without considerable interruptions) for at least 80 days. The patient characteristics of the two groups are shown in Table 1. For the

adopters, usage of blood glucose, insulin, and food measurements was relatively frequent and stable. It is notable, as can be seen in Figure 2, that there was much variation in the usage among patients within the adopter group.

Periodicities

The uneven sampling rate of the SMBG measurements allows for period detection at frequencies above the Nyquist frequency by using periodograms for unevenly sampled data such as the Lomb–Scargle periodogram.[16,17] The Lomb–Scargle periodogram is a linear best fit to a sine curve with the given period and arbitrary phase. Using the fact that the power is exponentially distributed one can find the significance of the period. If a significant period is identified, this can be visualized to the user for example in the form of a smoothed kernel regression curve over the given period.[18]

We searched for periodicities in the complete data from all adopters by computing the Lomb–Scargle periodogram for the

TABLE 1. PATIENTS CHARACTERISTICS
FOR THE GROUPS OF ADOPTERS AND NONADOPTERS

Characteristic	All patients (n = 30)	Nonadopters (n = 12)	Adopters (n = 18)
Age years [mean (SD)]	39.1 (11.2)	32.9 (10.0)	43.2 (10.3)
Female (%)	13 (43.3)	4 (33.3)	9 (50)
HbA1c (%) [mean (SD)] at			
Start	8.3 (1.4)	8.7 (1.7)	8.1 (1.2)
End	8.3 (1.3)	8.9 (1.1)	8.1 (1.3)

HbA1c, glycated hemoglobin.

TABLE 2. ESTIMATED P VALUES FOR EACH PERIODICITY AND EACH PATIENT

P number	Daily	Weekly	Monthly
1	0.0057 (24 h)[b]	0.051 (185 h)	—
2	$<10^{-5}$ (24 h)[b]	$<10^{-5}$ (168 h)[b]	0.00015 (730 h)[b]
3	0.00063 (24 h)[b]	0.012 (149 h)[a]	—
6	$<10^{-5}$ (24 h)[b]	0.0033 (170 h)[b]	0.058 (614 h)
7	0.00040 (24 h)[b]	0.0057 (165 h)[a]	—
8	$<10^{-5}$ (24 h)[b]	0.0086 (166 h)[a]	—
10	0.00006 (23.6 h)[b]	0.039 (191 h)	0.090 (671 h)
12	$<10^{-5}$ (24 h)[b]	0.0061 (145 h)[a]	—
14	$<10^{-5}$ (24.2 h)[b]	0.0015 (165 h)[b]	0.0026 (670 h)[b]
15	$<10^{-5}$ (24 h)[b]	0.00089 (152 h)[b]	0.0048 (701 h)[a]
16	0.00018 (24 h)[b]	0.0070 (166 h)[a]	0.00078 (728 h)[b]
20	$<10^{-5}$ (24 h)[b]	0.0033 (144 h)[a]	0.019 (743 h)
22	$<10^{-5}$ (24.1 h)[b]	$<10^{-5}$ (153 h)[b]	$<10^{-5}$ (852 h)[b]
26	$<10^{-5}$ (23.9 h)[b]	0.00088 (177 h)[b]	0.00011 (645 h)[b]
28	$<10^{-5}$ (24 h)[b]	0.00024 (171 h)[b]	0.00092 (847 h)[b]
29	$<10^{-5}$ (24 h)[b]	0.041 (157 h)	—
30	$<10^{-5}$ (24.1 h)[b]	0.0010 (168 h)[b]	0.047 (705 h)

A dash indicates where no significant peak with $P < 0.1$ was found. Numbers in parentheses indicate the period at which the spectral maximum is, where we have searched within 24 ± 1 h, 168 ± 24 h, and 720 ± 150 h, respectively.
[a]$P < 0.05/3$, [b]$P < 0.01/3$.

period ranges 24 ± 1 h, 168 ± 24 h (weekly), and 720 ± 150 h (30 days/month), and we identified the most significant peak in each range. If the most significant peak occurs at a frequency f and has power $P(f)$, the P value is approximated as $p = e^{-P(f)}$. To confirm the validity of this assumption we performed Monte Carlo tests for a few selected values and found good agreement.

In Table 2, the significance of periodicities on a daily, weekly and monthly basis for each patient in the adopter group is shown. Because we used the same data to test for three different periodicities, we corrected the significance levels by Bonferroni's correction, such that a 0.05 overall significance level corresponds to a $0.05/3 = 0.017$ individual significance level. Thus, any patient who has a significant periodicity has consistent sine-like variation with that period. For instance, a patient's blood glucose level may be typically higher in the evenings, which will emerge as a significant periodicity after data from a certain number of days are recorded. The pattern can subsequently be visualized to the patient so that the patient can be informed and make appropriate changes or discuss the pattern with his or her physician.

For visualization of the periodic pattern, we use kernel regression smoothing with a fixed bandwidth. It is still necessary to display an indication of the error because this will typically vary. In particular, there are likely to be few measurements (and high error) during the night. The nightly measurements may also carry higher bias because patients are more likely to wake up and measure their blood glucose when it is outside the normal range.

Pattern recognition

Using and fine-tuning pattern recognition methods is laborious, but despite our best efforts we could not find any technique that performed substantially better than random at predicting future blood glucose values based on statistical learning techniques.[19] We also constructed a three-way clas-

sification problem for the classes low, high, and normal range, which is a simpler problem compared with predictive regression. The limits of the ranges of each class were determined individually in the training set by the 20% and 80% quantiles in order to avoid empty training sets. Thus, the classes did not necessarily reflect medical hypo- or hyperglycemia, but rather unusually low or high blood glucose values. The classifier was trained on the data recorded up to the current time and predicted the probability of low or high glucose at the next blood glucose measurement. If a pattern that indicated low blood glucose was identified, for example, the important predictors could be presented to the user, enabling the user to learn from typical undesirable situations that had not been identified previously.

The results indicate that pattern recognition methods for prediction or predictive classification are not useful for these data. We tested several different classifiers, including random forests, support vector machines, and quadratic discriminant analysis. We found that generally support vector machines with radial basis function performed the best of these, including providing the best balancing of the classes. However, the classification was not good enough to provide meaningful insight from the feature selection (i.e., we could not discriminate which features were important for classification). We performed two independent binary classifications for the problems low versus not low and high versus not high, and whenever there were conflicting results for any data point, the prediction was set to the normal range. Thus we could compute the precision and recall for the results, and the corresponding points in the receiver operating characteristics plot shown in Figure 3 for both problems. Although most results
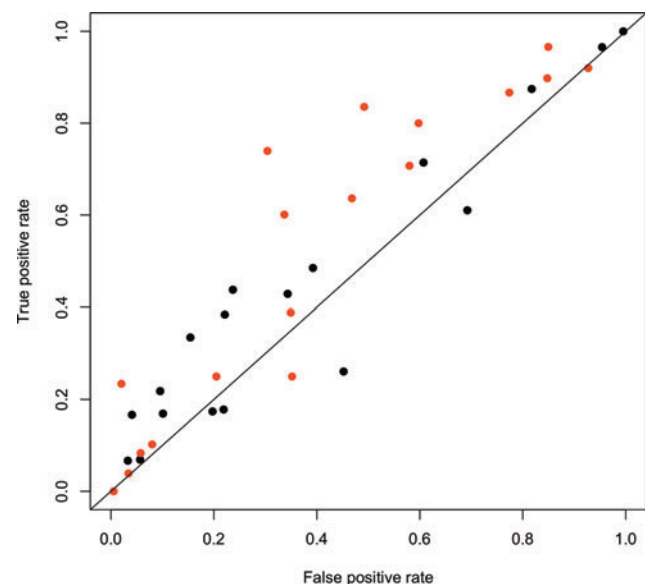

FIG. 3. Receiver operating characteristics plot for classification of low or high blood glucose using support vector machines with a radial kernel function. Gray dots indicate high glucose (80% quantile), black dots low glucose (20% quantile), and each dot corresponds to one patient. The diagonal line shows the random classifier, points closer to the upper left corner than the diagonal line mean improved classification over random guessing. Color graphics available online at www.liebertonline.com/dia
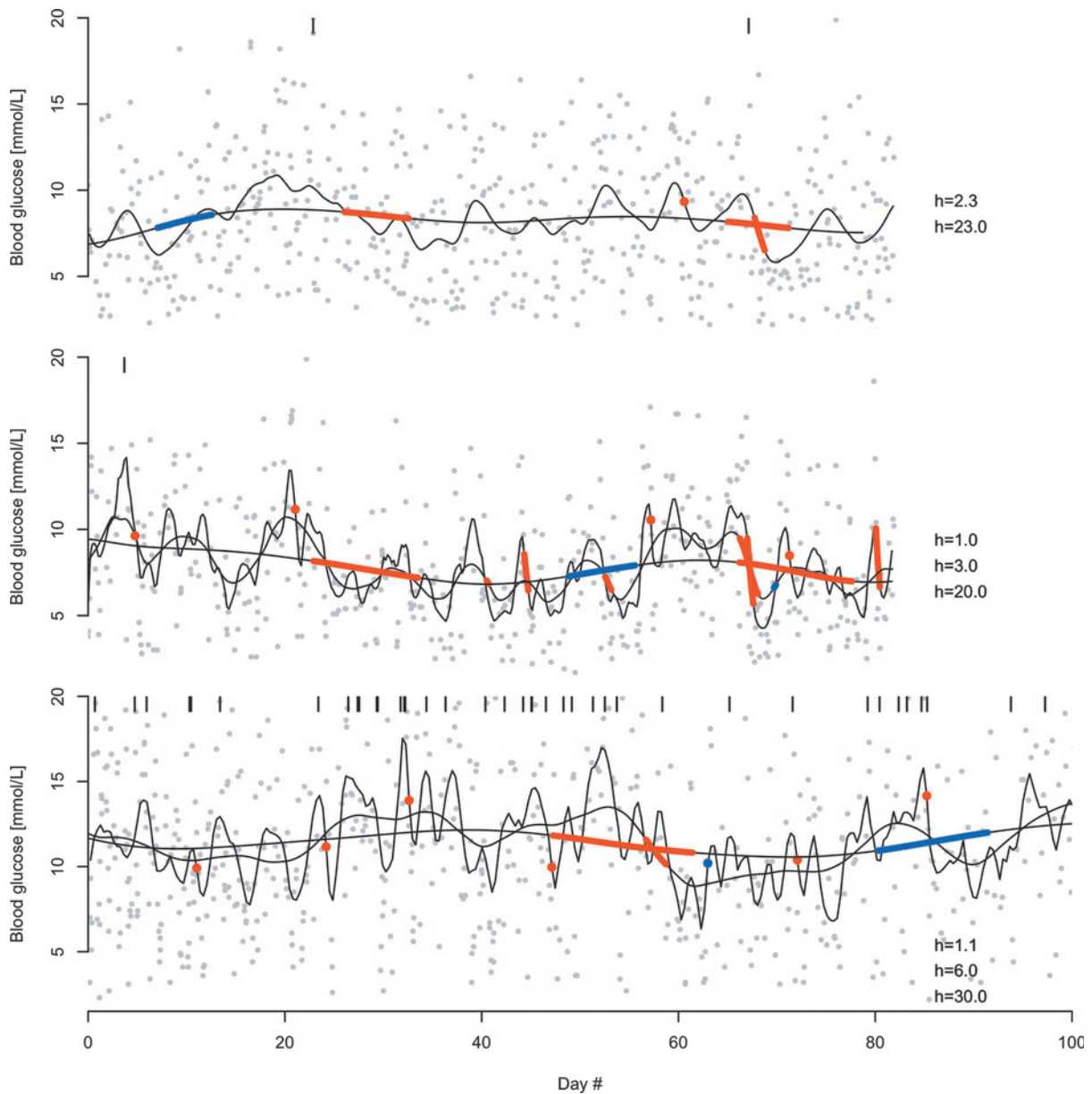
FIG. 4. Significant scale-space trends for three selected patients. Gray dots are blood glucose measurements, the *y*-axis is cropped for visibility, and black vertical lines indicate measurements above the selected range. Black smooth lines are kernel regression functions using linear least squares regression with a quartic kernel function and bandwidths in hours as indicated in the individual panels. Red segments show a significant decreasing trend at the given scale, and blue segments indicate a significant increasing trend. The scales shown are chosen based on the full c-SiZer plots such that most significant trends are shown, although some details on the very smallest scales are omitted for clarity. Bandwidth is here defined as half of the support of the kernel.

were above the non-discrimination diagonal line, they were only slightly better than random guessing. The best prediction results were for hyperglycemia, performing somewhat better than the best hypoglycemia results.

Significant change points

The large amount of noise in the SMBG measurements makes it difficult to identify trends apart from obvious ones when a simple scatterplot is presented such as in the current version of FTA (Fig. 1A). SMBG values are not necessarily an

unbiased sampling of actual blood glucose because measurements typically are preprandial or measured when the user suspects low or high blood glucose values. Nevertheless, trends in the measured values can be valuable as indicators of physiological change.

An appropriate tool to identify changes in live systems is the c-SiZer algorithm, which detects changes very early and at any scale of observation (S.O. Skrøvseth, J.G. Bellika, and F. Godtliebsen, "Causality in scale space as an approach to change detection," manuscript submitted for publication). The latter property is important because trends may appear at

distinct scales simultaneously. For example, a user may experience a short-term significant increase in blood glucose as part of a long-term significant decrease. c-SiZer allows for detection and visualization of these trends as they appear.

c-SiZer is based on the SiZer (Significant Zero-crossings of derivatives) methodology, which defines a scale space spanned by a bandwidth and time and performs a hypothesis test of the derivative of the signal at every point in scale space.[20] c-SiZer is a causal version that is adapted for live processes such as sensor data.[21] Thus changes can be detected live during data acquisition and presented to the user.

All adopters had significant trends at some point, most commonly on smaller scales, typically of one or a few hours. Large-scale trends on several days or more also appeared in most patients.

Example of three patients' time series and significant trends are shown in Figure 4. All these patients exhibit trends on both small and large scales. Although some trends are significant at several scales, in most cases a trend appears on only one of the selected scales. On a few occasions there are also conflicting trends (i.e., a significant increasing trend on one scale may overlap with a significant decreasing trend on another scale). There is no contradiction in this because long-term trends inevitably contain small-scale variations that may themselves be significant.

Discussion

We have presented three distinct methods for analysis and presentation of data recorded, primarily blood glucose measurements, by the FTA system. Even with a platform that is designed to be as nonintrusive as possible, the data quality is usable for some particular feedback algorithms.

The poor performance of the pattern recognition is likely to be largely caused by the low fidelity of the food recording. The system has a minimally functional food recording system in which the users themselves are free to define which foods go into which categories. Thus, the correspondence between carbohydrate content and food categorization is minimal, while all other nutritional information is absent. In future versions of the FTA we will experiment with implementing carbohydrate counting, analogous to recent commercial devices. Providing feedback on variability to patients could potentially be another useful feature, but given the uncertainties there is no unambiguous way to implement such features based on SMBG data.

The FTA itself has proven a valuable tool for T2D patients, and the current study shows that it is also useful for T1D patients, with the majority of the users actively using the device for a long period. Although some of the extended usage is likely to be explained as a Hawthorne or novelty effect, it is probable that each of these would wear off over the study period. We find little evidence of that given that usage patterns are consistent over the whole period for the active patients. In this study, the FTA was implemented on Windows Mobile 6.5 phones, which were already outdated at the point of the study. This was a common source of complaint among the users, particularly those accustomed to more modern devices or systems such as Android- or iOS-based devices. The application is currently being improved and developed on the Android system, and the effect of both platform and device on user satisfaction and adherence is being investigated.

For all applications a significant amount of training of the patients has to be performed. A user cannot be expected to understand the ideas immediately. Once minimal training has been performed and visualized feedback becomes available, the motivation to improve the quality of the recordings can be substantiated and thus initiate a positive feedback loop from the patient's perspective.

When trends and periodicities are presented to a user, the results will be provided in real time with only those properties that are currently relevant and significant visible to the user. Thus the display will be simplified and will not always have any pattern displayed.

Acknowledgments

Author Disclosure Statement

No competing financial interests exist.

References

1. Hovorka R: Closed-loop insulin delivery: from bench to clinical practice. Nat Rev Endocrinol 2011;7:385–395.
2. Dolgin E: Medical devices: Managed by machine. Nature 2012;485:S6–S8.
3. Johansen MD, Hejlesen OK, Cavan DA: Hypoglycemia impairs quality of blood glucose simulation in a clinical decision support system. J Diabetes Sci Technol 2011;5:894–900.
4. Pickup JC, Freeman SC, Sutton AJ: Glycaemic control in type 1 diabetes during real time continuous glucose monitoring compared with self monitoring of blood glucose: meta-analysis of randomised controlled trials using individual patient data. BMJ 2011;343:d3805.
5. Ceriello A, Ihnat MA: "Glycaemic variability": a new therapeutic challenge in diabetes and the critical care setting. Diabet Med 2010;27:862–867.
6. Siegelaar SE, Holleman F, Hoekstra JBL, DeVries JH: Glucose variability; does it matter? Endocr Rev 2010;31:171–182.
7. Dorr D, Bonner LM, Cohen AN, Shoai RS, Perrin R, Chaney E, Young AS: Informatics systems to promote improved care for chronic illness: a literature review. J Am Med Inform Assoc 2007;14:156–163.
8. Chomutare T, Fernandez-Luque L, Årsand E, Hartvigsen G: Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. J Med Internet Res 2011;13:e65.
9. Harris LT, Tufano J, Le T, Rees C, Lewis GA, Evert AB, Flowers J, Collins C, Hoath J, Hirsch IB, Goldberg HI, Ralston JD: Designing mobile support for glycemic control in

patients with diabetes. J Biomed Inform 2010;43(5 Suppl):S37–S40.

10. Liang X, Wang Q, Yang X, Cao J, Chen J, Mo X, Huang J, Wang L, Gu D: Effect of mobile phone intervention for diabetes on glycaemic control: a meta-analysis. Diabet Med 2011;28:455–463.

11. Charpentier G, Benhamou P-Y, Dardari D, Clergeot A, Franc S, Schaepelynck-Belicar P, Catargi B, Melki V, Chaillous L, Farret A, Bosson JL, Penfornis A; TeleDiab Study Group: The Diabeo software enabling individualized insulin dose adjustments combined with telemedicine support improves HbA1c in poorly controlled type 1 diabetic patients: a 6-month, randomized, open-label, parallel-group, multicenter trial (TeleDiab 1 Study). Diabetes Care 2011;34:533–539.

12. University Hospital of North Norway: Self-Management in Type 2 Diabetes Patients Using the Few Touch Application. http://clinicaltrials.gov/ct2/show/NCT01315756 (accessed March 22, 2012).

13. Årsand E, Tatara N, Østengen G, Hartvigsen G: Mobile phone-based self-management tools for type 2 diabetes: the Few Touch Application. J Diabetes Sci Technol 2010;4: 328–336.

14. Gammon D, Årsand E, Walseth OA, Andersson N, Jenssen M, Taylor T: Parent-child interaction using a mobile and wireless system for blood glucose monitoring. J Med Internet Res 2005;7:e57.

15. R Development Core Team: R: A Language and Environment for Statistical Computing. 2011. www.R-project.org (accessed September 18, 2012).

16. Scargle JD: Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data. Astrophys J 1982;263:835–853.

17. Skrøvseth SO, Godtliebsen F: Scale space methods for analysis of type 2 diabetes patients' blood glucose values. Comput Math Methods Med 2011;2011:1–7.

18. Wand MP, Jones MC: Kernel Smoothing. Boca Raton, FL: Chapman & Hall/CRC, 1995:212.

19. Hastie T, Tibshirani R, Friedman JH: The Elements of Statistical Learning. New York: Springer-Verlag, 2008:745.

20. Chaudhuri P, Marron JS: SiZer for exploration of structures in curves. J Am Stat Assoc 1999;94:807–823.

21. Skrøvseth SO, Dias A, Gorzelniak L, Godtliebsen F, Horsch A: Scale-space methods for live processing of sensor data. Stud Health Technol Inform 2012;180:138–142.

Address correspondence to:
Stein Olav Skrøvseth, Ph.D.
University Hospital of North Norway
Norwegian Centre for Integrated Care and Telemedicine
P.O. Box 35
9038 Tromsø, Norway

E-mail: stein.olav.skrovseth@telemed.no