Estimating Blood Sugar from Voice Samples

A Preliminary Study

C. Tschöpe, F. Duckhorn
Fraunhofer IKTS Dresden
Dresden, Germany

M. Wolff, *Member, IEEE*
BTU Cottbus-Senftenberg
Cottbus, Germany

G. Saeltzer
Dresden, Germany

Abstract—**Diabetes is a widespread disease. Patients have to pay attention to their blood sugar value all the time. Hypoglycemia and hyperglycemia are dangerous and must be treated. Self-testing, typically involving a prick with a lancet, is often painful and the costs for testing devices and supplies are immense. It is well-known that the human voice carries all kinds of information and depends on various factors. Therefore we conjectured that it may also be influenced by the blood sugar level. This paper presents a preliminary study which shows that a patient's blood sugar condition actually seems to manifest itself in the voice. Results encourage us to go the next steps: a large-scale long time study in collaboration with medical experts and, finally, the development of a measurement device or smartphone app.**

Index Terms—**diabetes, blood sugar value, voice characteristics, voice analysis**

I. INTRODUCTION

Diabetes is a widespread disease which concerns more and more people. According to the World Health Organization (WHO) "in 2014 the global prevalence of diabetes was estimated to be 9 % among adults aged 18+ years" [1]. Consequences can be fatal: "Diabetes is a leading cause of blindness, amputation and kidney failure" [1]. Hence it is very important for diabetes patients to monitor their blood sugar level up to five times a day. Common practice for self-control and before injecting insulin is to use lancets. Patients extract a small amount of blood and test their blood sugar level. This procedure is painful and nasty. If the device does not work correctly the patient has to repeat the procedure. In addition, taking blood increases the risk of infections if lancets or injection sites are not sterile. Besides the discomfort, patients and health systems spend a lot of money. For instance, yearly costs for measurement devices, test strips and lancets amount to 2-3 billion euros in Germany.[1]

II. HYPOTHESIS

This preliminary study aims to test the theory that the blood sugar condition manifests itself in the human voice. Hence by means of machine learning and pattern recognition algorithms at least a coarse automatic estimate like "hypoglycemic", "normal", or "hyperglycemic" state should be possible.

Our hypothesis is substantiated by experiences of "expert" listeners who have lived together with diabetes patients for a long time and are able to recognize the blood sugar condition from subtle cues in the familiar person's voice. It seems that the individual voice of a patient carries sufficient information on this medical condition.

This observation inspired us to apply results of our long-standing research in intelligent signal processing and pattern recognition [3] to a new application from medicine. In a former paper we already investigated auscultatory blood pressure measurement using acoustic pattern recognition [4].

III. EXPERIMENTAL SETUP

A. Datasets

We collected a small dataset of voice recordings and corresponding invasively measured blood sugar values from two individuals (person A: $M_A = 80$ measurements and person B: $M_B = 60$ measurements). To eliminate unnecessary variance we asked the patients only to speak the German sentence "Hallo, wie geht es Dir?" (Hello. How are you?) and to repeat that sentence several (2-6) times per invasive measurement. Blood sugar levels vary from 3.1 to 18.1 mmol/l for person A and from 5.7 to 13.0 mmol/l for person B and pass the Shapiro-Wilk test for normality after a mel-transform [5] with spectral wraping factor $\lambda = 0.25$:

$$y = f(x) = x + \frac{40}{\pi} \arctan \frac{\lambda \sin \frac{\pi}{20}x}{1 - \lambda \sin \frac{\pi}{20}x}. \tag{1}$$

Transformed features have a mean value $\mu_y = 6$ mmol/l and a standard deviation $\sigma_y = 2$ mmol/l.

B. Voice Feature Extraction

From the voice recordings we extracted 4368 different voice features using the openSMILE toolkit [6]. We excluded 1993 features which failed the Shapiro-Wilk test for normality. The remaining 2375 features were averaged over the repeated recordings for the same blood sugar level in order to reduce noise. Finally we sorted the averaged features by descending Pearson correlation with the invasively measured blood sugar value and selected the N best features ($10 \leq N < M_i$) for a correlation analysis. The N-best feature sets were chosen individually for persons A and B.

---

[1]Estimate by the authors from own experience based on a number of 6 million diabetes patients in Germany [2].

CPS
Conference Publishing Services

C. Automatic Estimation of the Blood Sugar Level

In order to test our hypothesis we performed a simple linear regression analysis with the individually selected N-best voice features as the independent variables and the invasively measured blood sugar level as the dependent variable (ground truth of pattern recognition).

As our datasets are very small and the feature extraction is very selective we compare the regression between the voice features and the actual blood sugar measurements with a second regression between the actual voice features and random blood sugar values. The latter were sampled from a normal distribution with $\mu = 6\,\mathrm{mmol/l}$ and $\sigma = 2\,\mathrm{mmol/l}$ (see section III-A). To ensure a fair comparison we selected the N voice features with the highest Pearson correlation to the random blood sugar value, just as we did for the real measurements.

If our hypothesis holds we expect the coefficients of determination (R^2) of the regression to be significantly higher for the real blood sugar measurements than for the random control samples.

IV. Experimental Results

Table I shows the coefficients of determination R^2 of the linear regression model for both test persons and for the real and random blood sugar values. As expected, the coefficients are higher for real blood sugar values than for random values. Statistical significance ($p < 0.05$ of null hypothesis $R = 0$) is attained only for the real blood sugar values, not for the random control set.

TABLE I
LINEAR REGRESSION RESULTS FOR BOTH TEST PERSONS
(NUMBER OF FEATURES: $N = 45$).

$N = 45$ Blood sugar	Person A		Person B	
	real	random	real	random
R^2	0.843	0.669	0.967	0.906
$p(H_0 : R = 0)$	0.000	0.133	0.001	0.124

To gain some more insight we varied the number N of selected voice features. As seen in Figs. 1 and 2, the coefficient for determination is higher for the real blood sugar values than for the random ones.

V. Conclusion

We conducted a preliminary experiment on a possible relation between the blood sugar level of patients and their voice quality. The first results are promising. At least for our small dataset of only two test subjects we were able to obtain linear regression models with coefficients of determination significantly different from zero. Comparison with a random control sample of blood sugar values indicates that our findings are unlikely to be attributed to chance. However, the specific voice features correlating to blood sugar level seem to be highly individual.

The experimental findings presented here seem to justify a large-scale study with many patients and in collaboration
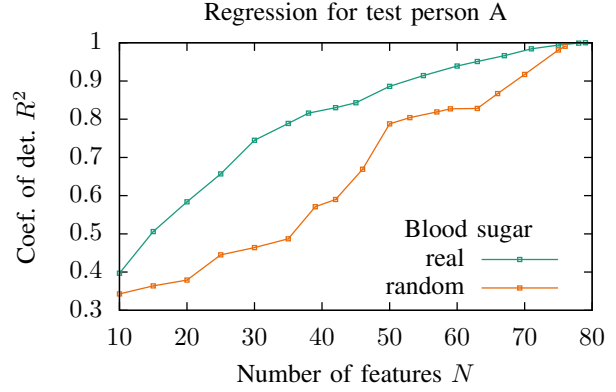

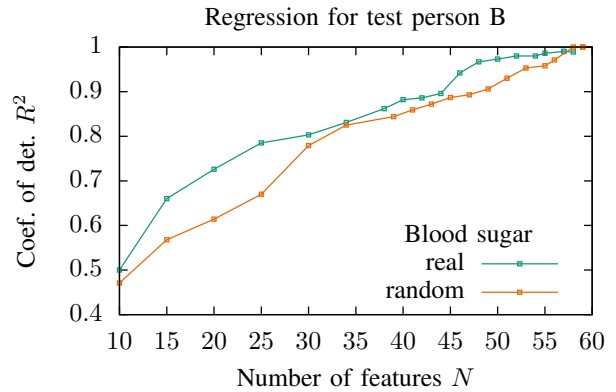Fig. 1. Variation of number of features for test person A.

Fig. 2. Variation of number of features for test person B.

with medical experts as a next step. The ultimate goal is a portable device or smartphone app which allows a noninvasive and comfortable check of the current blood sugar level for diabetes patients. Respective patents are pending.

VI. Acknowledgment

References

[1] *Global status report on noncommunicable diseases 2014*. Geneva, World Health Organization, 2012.

[2] T. Danne, E. G. Siegel, *Nationale Diabetes-Strategie: Alle Zeichen auf Grün!?*, Deutsche Diabetes-Hilfe: Deutscher Gesundheitsbericht Diabetes 2015, pp. 5–7.

[3] C. Tschöpe, M. Wolff, *Statistical Classifiers for Structural Health Monitoring*, IEEE Sensors Journal, Volume 9, No. 11, Nov. 2009, pp. 1567–1576.

[4] M. Wolff, U. Kordon, H. Hussein, M. Eichner, R. Hoffmann, C. Tschöpe, *Auscultatory Blood Pressure Measurement using HMMs*, IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2007), Honolulu, Hawaii, 15-20 April 2007, in Proc. Vol. 1, pp. 405–408.

[5] G. Strecha, *Skalierbare akustische Synthese für konkatenative Sprachsynthesesysteme*, in Studientexte zur Sprachkommunikation, Vol. 77, TUDpress, Dresden, Germany, 2014.

[6] F. Eyben, F. Weninger, F. Gross, B. Schuller, *Recent Developments in openSMILE, the Munich Open-Source Multimedia Feature Extractor*, in Proc. ACM Multimedia (MM), Barcelona, Spain, ACM, ISBN 978-1-4503-2404-5, pp. 835–838, October 2013.