

<https://doi.org/10.1038/s43856-025-01241-7>

Clinical validation of noninvasive blood glucose measurements by midinfrared spectroscopy

Check for updates

Michael Kaluza¹, Sergius Janik¹, Thorsten Lubinski¹, Mattia Saita ¹, Peter Lachmann ¹, Luca Canini¹, Valentino Lepro ¹ & Werner Mäntele ^{1,2}

Abstract

Background Non-invasive glucose monitoring (NIGM) is considered the holy grail of diabetes technology. Currently, diabetes is predominantly managed based on finger pricking and analysis of a drop of blood with test strips fitting into a glucometer. This invasive, painful and uncomfortable procedure is one of the reasons for insufficient diabetes management. Sensors for continuous glucose monitoring (CGM) in interstitial fluid (ISF) using a subcutaneous microscopic filament present a minimally invasive alternative. Worldwide, there is intense research and development for a true non-invasive glucose measurement.

Methods DiaMonTech has demonstrated a NIGM technology that targets glucose in ISF. An infrared beam from a quantum cascade laser excites glucose molecules at wavelengths between 8 and 12 μm , where glucose has specific fingerprint absorbance. Absorption results in a small amount of heat in the skin, which can be detected on the surface using a photothermal deflection technique. This procedure is painless, harmless, and does not require consumables.

Results We report here a single-center clinical test with 36 individuals (clinicaltrials.gov ID: NCT06088615). The accuracy of the NIGM device was evaluated in two subsequent sessions per individual, with different amounts of calibration data. Four different algorithms were tested for data analysis. The accuracy for the best algorithm, expressed as Mean Absolute Relative Difference was 20.7 % and 19.6 % for the two sessions.

Conclusions This is equivalent to the performance of early CGM systems cleared by the FDA for adjunctive use by people with diabetes. It demonstrates that glucose can be reliably measured with this non-invasive technology and opens new perspectives for a better management of diabetes.

Plain language summary

People with diabetes cannot regulate the amount of a sugar called glucose in their blood. They need to regularly check the amount of glucose in their blood so they can take treatments that adjust the levels. This procedure is painful and uncomfortable as it involves pricking the skin to make it bleed. We used an infrared laser that enables a painless determination of the amount of glucose present in skin. Our device was tested in 36 people with and without diabetes and showed an accuracy similar to other glucose monitors approved by the US food and drug administration (FDA). Our method could make blood glucose checks easier and more comfortable in the future and improve diabetes management.

Non-invasive glucose monitoring is a complex problem and considered the holy grail of diabetes technology, as it can improve the life of many humans around the globe. An increasing number of people (currently 589 million adults) worldwide live with diabetes¹. At present, diabetes cannot be cured, instead it is managed by strict control of blood glucose (BG) and adaptation of food intake, physical activity, medication, and if necessary, by insulin administration. BG monitoring is predominantly performed invasively by finger pricking and analysis of a drop of blood using a glucometer in combination with enzymatic test strips. This

procedure produces problematic waste², bears the risk of infection, and is uncomfortable and painful. Consequently, users of traditional devices do not perform the close-meshed monitoring needed to keep their BG in the “green range” (approx. 70–180 mg/dL). The situation has considerably improved by the introduction of minimally invasive sensors for continuous glucose monitoring in the interstitial fluid (ISF), where a short filament is introduced into the skin using a needle (for a review, see ref. 3). These sensors are particularly appreciated by people with type 1 diabetes who require close-meshed control with warnings for hypo- and

¹DiaMonTech AG, Berlin, Germany. ²Institut für Biophysik, Johann Wolfgang Goethe-Universität, Frankfurt am Main, Germany.

e-mail: werner.maentele@diamonotech.de; maentele@biophysik.uni-frankfurt.de

hyperglycemia, although skin irritations (the sensor patch is worn for up to two weeks) have been reported⁴.

In the search for a truly non-invasive solution, numerous technological attempts have been made over the past >20 years; however, no reliable glucometer function has been demonstrated in a consumer device. Some of these efforts included physiological parameters as proxies that parallel BG, such as temperature, electric conductivity, impedance of skin, or ultrasound and microwave transmission of skin (for a recent review, see ref. 5). The specificity of these proxies for BG concentrations, however, is too low to allow a safe measurement for people with diabetes. Other attempts use optical techniques such as near-infrared spectroscopy (for a recent review, see ref. 6), Raman spectroscopy⁷⁻⁹, or mid-infrared spectroscopy¹⁰⁻¹³ directly targeting the glucose molecule.

Several approaches included the analysis of easily accessible body fluids such as urine, saliva, sweat and tear fluid as potential proxies for blood. Although these proxy fluids may contain some glucose, its concentration is much lower than in blood, it does not exhibit an explicit relation to BG, or it exhibits delays with respect to BG that could pose a health risk for people with diabetes.

ISF has turned out to be the most suitable proxy for blood. It is found in skin layers immediately below the *stratum corneum* at a depth from 20 to 200 μm below the skin surface and thus accessible with many optical technologies. It typically contains 85–90% of BG, equilibrated with the capillary blood flow by diffusion. Depending on blood circulation and local capillarization at the respective skin parts, it follows BG with delays of only a few minutes¹⁴. ISF contains essentially the blood protein albumin, glucose, lactate and ions¹⁰. CGM systems that measure with a small filament under the skin demonstrate that ISF is a reliable proxy for blood.

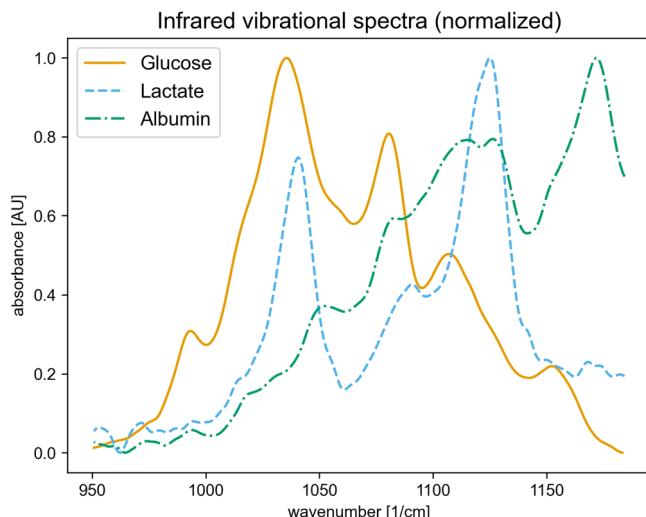


Fig. 1 | Infrared vibrational spectra of molecules relevant for skin in aqueous model solutions: glucose, lactate and albumin.

Optical technologies for NIGM targeting glucose in skin have been developed based on the absorption of visible, near-infrared (NIR), mid-infrared (MIR), Terahertz (THz) and microwave (MW) radiation as well as on Raman scattering. A glucose solution in water is colorless, i.e., visible light does not exhibit any specific absorption for the glucose molecule. Absorption of a glucose solution in the ultraviolet region (UV) is not specific, as is its absorbance in the THz or MW range. However, glucose exhibits a distinct vibrational signature in the MIR wavelength range from ~8 to 11 μm , arising from coupled -C-O- stretching and -O-H bending vibrations of the molecule that also are the basis of its Raman scattering behavior (Fig. 1). This signature is highly specific for the glucose molecule and allows distinction from other molecules in the skin but requires more sophisticated light sources, optics and detection. As for many molecules, the fundamental vibrations in the MIR are accompanied by overtones ("harmonics"), multiples of the fundamental modes that are found in the NIR, in the range between approx. 1 to 3 μm . The absorption cross-sections of these overtones are much weaker (more than 10–100 times) than those of the fundamentals. In addition, overtones of all molecules in skin are found in a relatively narrow spectral region, and furthermore, are affected by temperature in position and intensity. In spite of all these drawbacks, the easy availability of NIR light sources (e.g., Telecom lasers), optics, and detectors have led to many attempts to develop NIR-based glucose sensing.

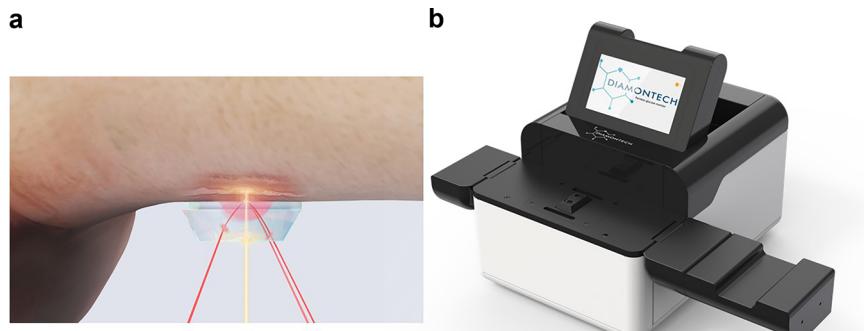
The outstanding MIR signature of glucose (Fig. 1) has prompted DiaMonTech to develop a sensor based on recently developed MIR sources, called quantum cascade lasers (QCL), and photothermal detection¹⁵⁻¹⁷. A block diagram of the working principle and the signal generation is provided in the section "Materials and Methods".

Briefly, pulsed MIR light from the QCL is directed through an internal reflection element (IRE) made from IR transparent material into the skin, where it penetrates up to 100 μm and is absorbed in ISF-containing layers (Fig. 2a). Glucose in these layers is vibrationally excited from the ground level v_0 to the first vibrational level v_1 , immediately (within $\sim 10^{-12}$ s) followed by thermal relaxation. The relaxation results in the deposition of a small amount of heat around the glucose molecule. Its amount corresponds to the energy of the absorbed MIR photon.

The heat spreads out to the surface of the skin and enters the IRE, where it forms a temporary thermal lens, a region where the refractive index is altered. To probe the thermal lens, a second visible or near-infrared laser beam is directed into the IRE and passes through the thermal lens, where it is deflected. The deflection is detected by a position-sensitive photo detector. The probe beam deflection is directly related to the pump laser power and the glucose concentration. Variation of the pump laser wavelength in the spectral region of glucose absorption allows to probe the glucose spectrum and to determine glucose concentration even on the background of other molecules in skin.

Based on this technology, DiaMonTech has developed prototypes of a table top glucometer ("D-Base") that include a commercial external cavity-tuned QCL, photothermal sensing unit, data acquisition unit and signal processing hardware and software (Fig. 2b).

Fig. 2 | Measurement principle and test device. **a** Visualization of photothermal detection of glucose in skin: In yellow the MIR excitation beam; in red the probe beam being deflected by the thermal lens. **b** D-Base with open display and mechanical lid (arm rest).



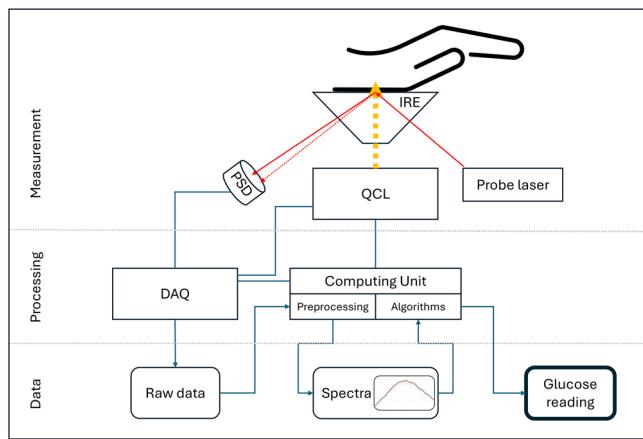


Fig. 3 | Mid-infrared measurement of glucose in skin. Block diagram of the Working principle and main components of the D-Base.

We report here a single-center clinical evaluation with 36 individuals. The accuracy of the NIGM device was evaluated in two subsequent sessions per individual, with different amounts of calibration data. Four different algorithms were tested for data analysis. The accuracy for the best algorithm, expressed as Mean Absolute Relative Difference, was 20.7% and 19.6% for the two sessions. This is equivalent to the performance of early CGM systems cleared by the FDA for adjunctive use by people with diabetes. It demonstrates that glucose can be reliably measured with this non-invasive technology and opens new perspectives for a better management of diabetes.

Materials and methods

The test device D-Base

The measurement principle has been described earlier¹⁶, has essentially remained unchanged and is illustrated in Fig. 3. The device has been improved, and the main changes are reported here.

The D-Base used in the study is the version “D-Base 2.0 prototype B”. Central component is a commercially available external cavity tunable quantum cascade laser (EC-QCL) as MID-IR source, the DRS Daylight Solutions Hedgehog (Leonardo DRS, San Diego, USA). This QCL is capable of emitting light in the 8.3–10.5 μm wavelength range and is used to excite the glucose molecules in the skin. The laser is operated in pulsed mode with 100 ns pulses at 10% duty cycle. Additionally, the QCL is controlled by a proprietary signal generator board that introduces an additional modulation of the laser, allowing emission only for 50% of the time with a frequency of 57 Hz while sweeping through the spectral range. The average output power of the QCL in this configuration is around 10 mW. This EC-QCL is classified as a Class 3R laser. The accessible emission limit (AEL) for the QCL based on the parameters used is 5000 W/m², while the measured value by an accredited testing site was 1373 W/m² of irradiance.

The probe laser to read out the thermal lens is used in total internal reflection in the IRE, thus only a few μW of power are accessible to the user. The AEL for the probe laser is 390 μW , while the measured value was below 10 μW . Both lasers are safely protected in the device and precautions have been taken to avoid the beam being viewed directly or through external optics. The device has passed the laser safety tests, which are based on the MPE (Maximum Permissible Exposure, see: IEC 60825-1 regarding laser safety) for skin. Our device works below 50% of the maximum safety threshold.

The optical measurement interface is the so-called internal reflection element (IRE), a ZnS crystal with antireflective coating manufactured specifically for this application. This ZnS crystal in contact with skin is insoluble and non-toxic and does not present any risk for the user. The probe laser deflection caused by the thermal lens is sensed by a proprietary custom-built detection system capable of resolving sub- μm movements of the probe laser,

and it is based on a 2D position-sensitive photodetector from Hamamatsu (Hamamatsu Photonics KK, Japan).

The raw data processing is handled by a custom-built data acquisition system (DAQ). The computing unit is based on a Raspberry Pi 4 (Raspberry Pi Holdings, Cambridge, England) with a 64 bit Linux OS and custom-built software.

A custom-made ergonomic wrist-rest accessory was designed and 3D printed. The object is designed to resemble a wristwatch and fastened with a Velcro strip to the arm. A mechanism allows the wristband to be securely and firmly clicked to the D-Base when the subject is ready for the measurements. This system ideally allows to measure the same skin spot over the entire session, improving measurement stability and reproducibility.

Study setup

The study reported here was performed between November 2023 and April 2024 by an independent institute (Institut für Diabetes-Technologie, IfDT) in Ulm (Germany) as a prospective study with 36 individuals (19 female, 17 male) aged between 20 and 75 years. The aim of this study was to assess the accuracy of this NIGM approach and to collect data for further product developments. The participants took part in three sessions on three different days. The first session was used for calibration purposes only, while the second session served as test session using only the first session for calibration. In the third session, the device was validated using the previous two sessions for calibration.

The study was planned in the summer of 2023 and the study design approved by the ethics committee and the competent authority. It was designed in a typical fashion for assessing the accuracy of a glucose measurement technology with an observation time of about 5 h per session and including an induced glucose change with a standardized meal and regular measurements with the reference and the test method. The study was approved by the Ethics Commission of the Landesärztekammer Baden-Württemberg on September 19, 2023, and registered by the BfArM on October 5, 2023, as well as at clinicaltrials.gov (ID: NCT06088615). It was conducted by the IfDT, an independent contract research organization accredited by the Deutsche Akkreditierungsstelle GmbH (DAkkS) as testing laboratory according to DIN EN ISO/IEC 17025 in terms of several test procedures according to DIN EN ISO 15197:2015.

In total, 36 participants were planned, recruited and enrolled for this study. A signed informed consent to participate in the study, including an understanding and willingness to follow the protocol, was required. Volunteers comprised people with diabetes (13 type-1 diabetes, 17 type-2 diabetes) as well as healthy individuals (6). One participant discontinued the study during the second session due to illness unrelated to the study, which led to 35 participants in total. One participant could only complete two of the three sessions due to a medical appointment unrelated to the study, hence a total number of 105 sessions were completed.

Important inclusion criteria for the test participants were clinical diagnosis of type-1 diabetes, type-2 diabetes, or no diabetes. For subjects with diabetes HbA1c < 10% was required. All participants had to be at least 18 years old, all genders could be included. All subjects were of white ethnicity.

No selection was made for skin pigmentation. Melanin (responsible for skin pigmentation) is produced by melanocytes in the deepest layer of the epidermis (“basal layer”) and stored there in intracellular units called melanosomes. This melanin distribution is applicable to all skin colors. Since the optical penetration depth of the infrared laser is only about 100 μm , no significant contribution from skin pigmentation is expected, which is a substantial advantage of our infrared technology.

The inner side of the distal forearm, preferably close to the wrist was selected as the measuring site, mainly because of the relatively thin stratum corneum (15–30 μm) and the fact that this body part is little affected by manual labor or exercise. Furthermore, this site was used because of its relevance for the development of wearable versions of the device. The subjects wore a wristband reproducibly attachable to the D-Base for the

duration of the session to improve data quality by standardizing the measurement spot and pressure.

Before the start of the session, the wristband was fixed to the inner side of the distal forearm, preferably close to the wrist of the participant, where it remained for the duration of the session (Fig. 4). In some cases, the wristband was placed slightly further up the inner side of the distal forearm to improve skin contact with the measurement interface. In a few cases, the wristband had to be moved slightly during the session to improve skin contact with the measurement interface. In both scenarios, no statistically significant impact on the device performance was observed.

Each measurement consisted of a reference measurement and a NIGM measurement, in this order. To further analyze the variations of raw spectra on identical measurement spots, the measurements were organized in sets of four with a droplet of water as contact medium. The participant placed the wrist on the device and did not lift it for four consecutive measurements. Each measurement took about 4 min, and measurements were performed in succession, so a set of four lasted approx. 20 min. In between these sets, the spectrum of a drop of water was measured to monitor device stability. The measurement time of approximately 4 minutes in this study was chosen for additional data collection and further analysis.

For each session, the participants arrived fasting at the study site in the morning. For the first hour, eight measurements were collected with fasting level glucose. Within the next four hours, up to 28 more measurements were collected with changing glucose levels induced by a standardized breakfast at varying times within the session.

The reference values of the first session were known and used for calibration purposes. For the second and third session, the reference data points were blinded by the study site until all respective NIGM measurement results were provided for the session to the study site. Four proprietary machine-learning algorithms were tested for the determination of glucose values, termed 1a, 1b, 2a and 2b. The main difference between the “1” and “2” algorithms were their complexity and the available parameter space. The difference between “a” and “b” algorithms was the recalibration with the first two measurement points of the test session and the glucose value span during fasting glucose measurements for the “b” types. Consistently, these two values were not used to determine the accuracy of the “b” algorithms. The results obtained by the NIGM algorithms were delivered once the respective session was completed. After the delivery of the NIGM measurement results was verified by the study site, the reference values were transferred via email to DiaMonTech. This process was quality-assured with a four-eye principle by the study site to exclude any data leakage.

In general, for the non-invasive measurement no significant discomfort or impact on tissue was reported by the subjects. This is to be expected as the tissue gets only warmed up below the user’s perception threshold at the laser spot. The non-invasive device has passed the relevant laser safety tests which are based on the MPE (Maximum Permissible Exposure) for skin. Our device works below 50% of the maximum safety threshold. Overall, ten adverse events (AEs) were recorded. Five AEs happened during the study participation without any relationship to the study or the device. The other five events were causally or possibly related to the

study procedures or the D-Base and all rated as mild. All AEs were rated as non-serious and were resolved within the follow up period. No corrective actions were needed.

Signal processing

The four machine-learning algorithms applied share the same pre-processing pipeline, i.e., the steps that transform the acquired raw data into features usable by the algorithms. Such pre-processing steps can be divided into two families: The first includes stateless steps that are applied independently to each individual measurement. The second is stateful and requires a calibration for each subject. Stateless methods include smoothing in the wavenumber domain by a Savitzky-Golay filter¹⁸ and unit normalization of measurements. This emphasizes relative amplitudes and spectral patterns over absolute amplitudes of features. Processing steps are applied in the following order: Smoothing, then unit normalization, followed by outlier detection and a Net Analyte Signal (NAS) algorithm (see below).

Stateful methods are used for two purposes: Outlier detection and discounting of non-glucose variations. A custom outlier detection algorithm was developed, based on unsupervised embedding learning and subsequent Hotelling’s T-squared-statistics analysis of the spectra’s semantic representation¹⁹. It is calibrated for each subject based on the first session (for the second session) or on the first two sessions (for the third session) and used to support the identification and handling of outliers (“bad measurements”) in subsequent sessions. Non-glucose variations are due to several factors, mainly caused by environment, instrument, and sample. They all vary across sessions, and it is impractical to model them individually. Our approach is to use a NAS unsupervised learning technique that is trained for each subject using the fasting level measurements available up to the session to be predicted²⁰. NAS transforms each measurement and presents the machine learning algorithms with features that highlight glucose-specific information.

Algorithms

The four machine learning algorithms are all based on the Support Vector Regression (SVR) algorithm^{21,22}. Compared to state-of-the-art neural networks, it is well-suited for working with scarce data. We preferred the SVR algorithm to the de facto standard in spectroscopy, PLSR, because of its ability to focus on hard-to-model measurements by targeting them as support vectors. In addition, the choice between different kernels during training allows further tuning for each subject. Model calibration for the second session is done on the first session, while the third session uses the first two sessions. Note that the first eight measurements of each session are not used to calibrate models, nor are they used to determine the accuracy: Rather, they are used for the unsupervised training of the NAS model.

The four algorithms were named 1a, 1b, 2a and 2b. “1” algorithms are limited to a linear SVR kernel because the physical phenomena under study are linear in nature. “2” algorithms, on the other hand, extend the kernel choice to better model the complexity of the real world. In addition, “b” algorithms, as opposed to “a” algorithms, employ a recalibration technique that adds the first two points of the test session and the fasting glucose



Fig. 4 | Photographs showing the use of a wristband to ensure reproducible attachment to the D-Base. Left: open wristband; center: closed wristband; right: wristband attached to D-Base.

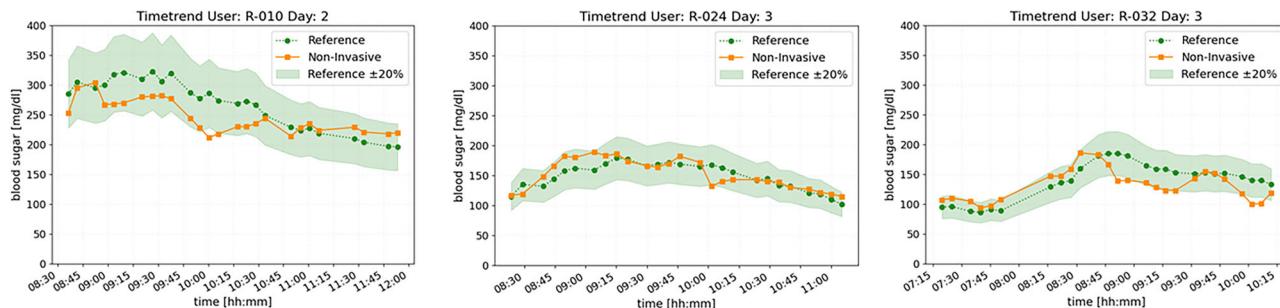


Fig. 5 | Time course of blood/tissue glucose for three individuals (R-010, R-024, R-032) at visits 2 and 3. Subject R-010 was a type-2 diabetes patient, subject R-024 a type-1 diabetes patient, and subject R-032 a healthy volunteer. Sample sizes were 26

measurements for each individual. The source data for Fig. 5 can be found in Supplementary Data

change to the calibration. This technique exposes the algorithm to measurements taken in the same environment, instrument and sample conditions of the measurements to be determined, without revealing much of the glucose test behavior, given its rather flat profile at the beginning. In total, up to 38 (74) points were available for calibration purposes per user for the second session (third session), while testing is computed on the up to 28 postprandial points per user. “b” algorithms, that use recalibration, have two points less for testing that become two more points for calibration. Combining the results of all users, the total number of points available for testing of the most promising algorithm 2b was 896 (862) for the second session (third session). One user only performed one test session. All available test points are included in the statistical analysis and the consensus error grid.

Each predictive model is calibrated via cross-validation over the corresponding training set. As a result, the best parameters across multiple predictive models might vary. The features used by the predictive algorithms are solely the acquired absorbance spectra; no measurement metadata (subject age, diabetes type, etc.) was used at this stage, as such information was disclosed to DiaMonTech only once the study was concluded.

The calibration is defined as “offline”, since algorithms to be tested on the second session are trained using data from the first session, while algorithms to be tested on the third session have the first and second session available as training set. This contrasts with “online” training, where an algorithm is trained incrementally as new data becomes available.

Over- and underfitting are relevant issues in the context of data scarcity and complex environments like non-invasive glucose measurements. We have mitigated this risk using several methods: To prevent overfitting, we use calibration datasets that are representative of the predictive problem, as the model sees at least a full glucose tolerance test with the glucose fluctuating. Our model choice, SVR, is less prone to overfitting²² compared to other popular models (e.g., random forests, which have been tested in the field by other research groups) and requires less data to produce models which are not underfitted, compared to models such as neural networks.

Reference device Contour Next

As a reference device, the Contour Next with test strips and control solution (Ascensia Diabetes Care Holdings AG, Basel, Switzerland) was used in this study in combination with the lancing device Medisafe Finetouch II (Terumo Corporation, Tokyo, Japan), measuring the capillary BG. Although not an official clinical reference method like venous BG analyzers, both are commercially available devices with well characterized performance and are considered representative for the standard measurement of a person with diabetes and sufficiently accurate as reference method in this study. We refrained from using a minimally invasive CGM sensor as a reference because of the lower precision and the potential delays with respect to blood glucose. BG measurements with Contour Next were performed in duplicate, and the average of the duplicate was used as reference value for the NIGM device.

No further medicinal products, human or animal tissues or their derivatives, or other biologically active substances were used.

Statistics and reproducibility

Consensus error grid. The consensus error grid (CEG) is a risk-oriented graphical representation of the test results²³. Glucose values are represented by the invasive reference value on the horizontal axis and the non-invasively measured value on the vertical axis. For an ideal measuring method, all data points would be found on the bold diagonal line. In reality, data points scatter along that line. The plot is divided by risk into zones. Zones A and B are considered “safe” for the user. Any clinically accepted invasive test strip device needs to perform with 99% of all measured data in zones A and B, among other requirements (ISO 15197:2013).

Bland–Altman plot. A Bland–Altman plot is an established method to assess the agreement between two measurement systems. It helps to visualize and analyze effects of a fixed or proportional bias. It shows the range within which ~95% of individual relative differences from the mean of test and reference system are found.

Mean Absolute Relative Difference. The Mean Absolute Relative Difference (MARD) is one of the de facto standard performance indicators of glucometers. It compares the measurement values of a test device with those of a state-of-the-art reference device and is an important tool for the comparison of different methods, technologies and products in comparable test settings. For each value pair the Absolute Relative Difference (ARD) is calculated and then the arithmetic mean of the ARD of all value pairs is used to assess the overall performance in a test session. We have thus decided to use the MARD score for the estimation of the accuracy instead of other metrics like RMSE or MAD. This is common in the comparison of different glucometer technologies and broadly used in specialized journals.

Results

Time course of blood/ISF glucose

Figure 5 exemplarily depicts for three participants the time course of BG measured with the invasive reference device and with the non-invasive technology D-Base using algorithm 2b. The shaded area around the reference values represents an absolute relative difference of 20% compared to the state-of-the-art reference method. The D-Base measurements follow closely the general shape (absolute values and trend) of the reference curve and are well within the acceptable band.

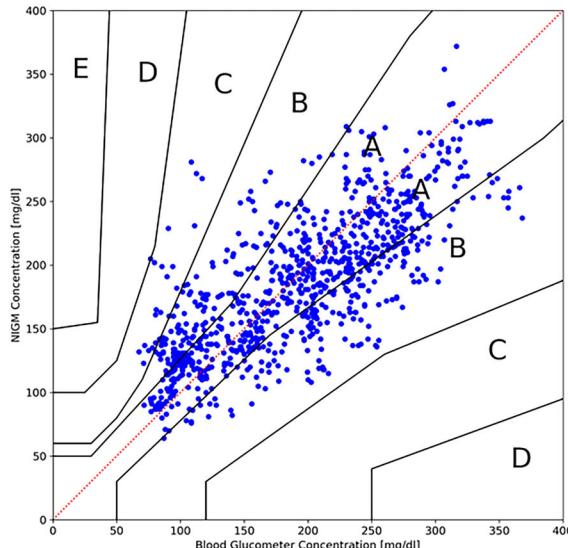
The best MARD values of 20.7% and 19.6% on the second and third sessions, respectively, were both obtained by algorithm 2b. The range of MARD values of the other algorithms was 22.9% to 35.3% (24.4% to 43.9%) for the second session (third session).

As for the performance separated by diabetes type, Table 1 summarizes the MARD values (average, 95% confidence interval) for the best performing algorithm 2b for the sessions 2 and 3.

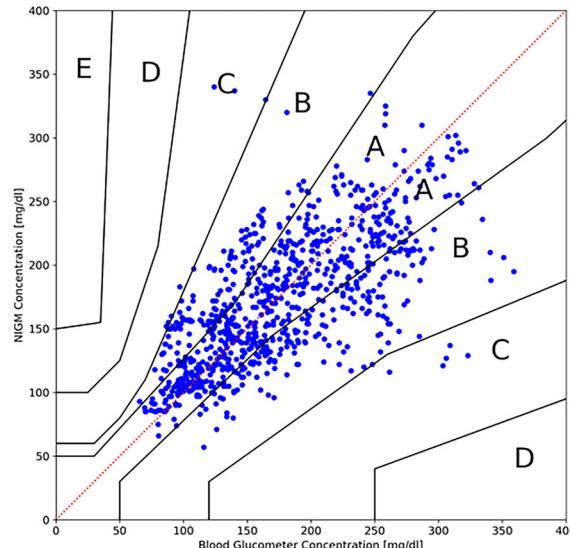
Table 1 | Performance of the non-invasive glucose measurement

Diabetes Type and Study Session	Global	Healthy	Type 1 diabetes	Type 2 diabetes
Session 2 MARD [CI 95%]	20.7% [16.7%, 24.5%]	15.8% [9.9%, 21.7%]	19.7% [13.4%, 26.0%]	23.2% [16.1%, 30.3%]
Session 3 MARD [CI 95%]	19.6% [16.6%, 22.9%]	12.7% [9.0%, 16.5%]	23.7% [16.3%, 31.1%]	19.4% [16.1%, 22.7%]

Consensus Error Grid (Session 2)



Consensus Error Grid (Session 3)

**Fig. 6 | Consensus error diagrams of non-invasive glucose measurements obtained with algorithm 2b for the second (left) and third session (right). Sample size was 896 (session 2) and 863 (session 3). The source data for Fig. 6 can be found in Supplementary Data****Table 2 | Range of reference BG values in the test sessions, globally and separated by diabetes type**

Reference BG	Global	Healthy	Type 1 diabetes	Type 2 diabetes
Min	66 mg/dL	66 mg/dL	68 mg/dL	75 mg/dL
Max	368 mg/dL	202 mg/dL	368 mg/dL	323 mg/dL

Clinical accuracy of the non-invasive method. The CEG allows a graphical representation of the test results, showing the comparison between invasive glucometer reference values and the NIGM measurements for all participants of the study. Figure 6 depicts the CEG of D-Base vs. Contour Next for the second and third session of algorithm 2b. The data includes a wide range of reference BG values from 66 mg/dL to 368 mg/dL, as summarized in Table 2.

For algorithm 2b, 97.5% (98.6%) of the values are within clinically accurate zones A + B in the second session (third session). Only one measurement fell into zone D, none in zone E. This performance is close to one of the requirements for invasive blood glucometers according to ISO 15197:2013.

For the other three algorithms, values in A + B of the CEG range from 89.1% (83.8%) to 96.5% (97.9%) for the second session (third session), with 0% (0%) up to 1.1% (1.2%) in zone D and none in zone E. The performance of all algorithms indicates the base measurement function of the technology, independent of the algorithm choice.

Analytical accuracy of the non-invasive method. The Pearson correlation coefficient is an indicator for the measurement capabilities of a glucose monitor and a value > 0.5 was designated to indicate an ability to follow changes of BG concentrations. This boundary was set as a success criterion for this study and could be shown on both visits for three out of

the four algorithms. The best performance was obtained for algorithm 2b, where Pearson correlation coefficients of 0.771 and 0.690 were achieved for the second and third sessions, respectively.

The agreement rates between the reference device Contour Next and the NIGM system for algorithm 2b showed the best results with 80.4% (second session) and 81.3% (third session) for the intervals of $\pm 30\%$ or ± 30 mg/dL for glucose concentrations ≥ 100 mg/dL or < 100 mg/dL, respectively.

A third method to analyze the measurement quality of the NIGM was the Bland-Altman plot, which indicates systematic influences and is depicted in Fig. 7. Again, algorithm 2b showed the best results out of the four algorithms. The 95% limit of agreement was $\sim \pm 40\%$.

Discussion

For clinical validation of new measurement technologies, measurement results are compared with reference values from a state-of-the-art glucometer during times of changing glucose levels. To classify clinical validations of glucose measuring techniques, it is helpful to distinguish between two different approaches.

In this manuscript, we refer to approaches as “retrospective” when the reference values are available before the measurement results are provided. These studies usually divide their datasets into a calibration and a validation dataset with varying segmentation of points. It is a common approach in early feasibility and explorative studies.

In contrast, an approach is defined as “prospective” when the non-invasive measurement results are provided before knowing the reference values (following an initial calibration). This study type is common for more mature technologies, simulating real-life scenarios.

As the retrospective data scientific problem to be solved is simpler, such approaches tend to yield a better Mean Absolute Relative Difference (MARD) than prospective ones.

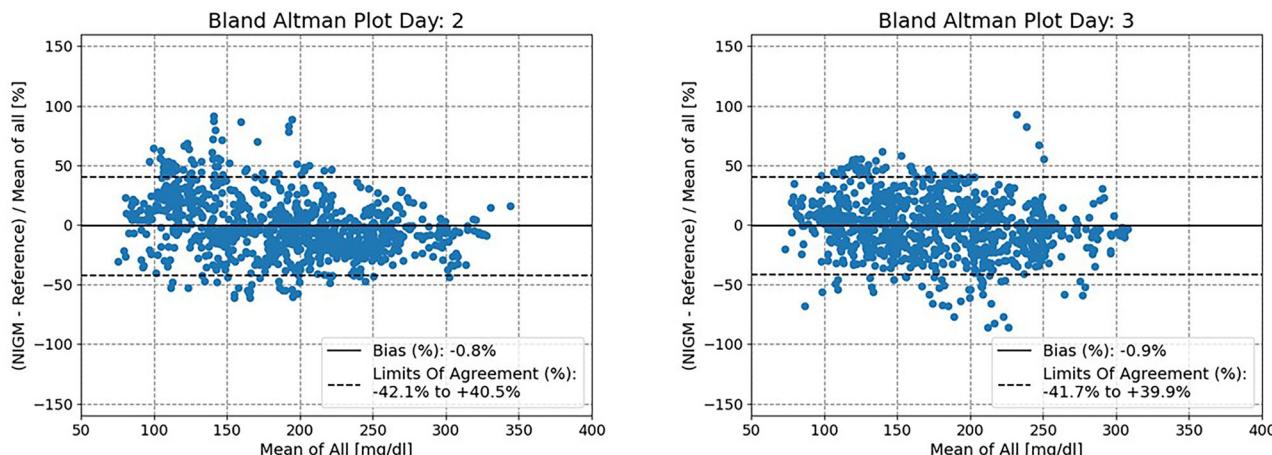


Fig. 7 | Bland–Altman plot for all measurements of the second (left) and third (right) session as obtained with algorithm 2b. Sample size was 896 (session 2) and 863 (session 3). The source data for Fig. 7 can be found in Supplementary Data

In a previous study (“Study 100”) with 100 volunteers, we demonstrated the feasibility of the technology. A MARD of 12% was achieved. That study has been done in a retrospective manner and indicated the feasibility of an NIGM approach based on MID-IR technology²⁴.

With the improved test device, the study we report here was conducted in a prospective fashion. For all utilized metrics, algorithms 2 showed better results than the corresponding algorithm 1, and algorithms b outperformed the respective algorithms a. Algorithm 2b delivered the best performance, closely followed by algorithm 2a. This is especially impressive, as the latter uses no recalibration from the session day.

The performance of algorithm 2b for the second and third sessions in MARD is comparable to early-stage CGM from Medtronic and Dexcom that the FDA approved for adjunctive use in 2005 and 2006, respectively^{25,26}. To our knowledge, no truly non-invasive glucometer has reached this MARD in a similar prospective clinical study setting.

In their early days, minimally invasive CGM had the disruptive advantage of augmenting, not replacing, state-of-the-art finger-prick measurements with a more frequent monitoring of blood sugar values. A truly non-invasive system promises a similar disruptive nature to the state-of-the-art methods of BG measurements with its ability of pain free and harmless use without consumables. Under the premise of a similar use as an adjunctive device without direct therapeutic consequence, it is hence reasonable to compare the performance to those early-stage CGM systems.

The performance of CGMs has improved over the years, so that current generations of commercially available CGM systems are cleared for non-adjunctive use. In a recent comparison of different invasive self-monitoring glucometers, MARD values from about 2.3% to about 20% were reported²⁷. For modern CGM systems, MARD values of up to 20% have been reported in an independent study²⁸, yet MARD values of around 10% have been published repeatedly in manufacturer-funded studies. As there are no relevant NIGM products established on the market, a similar scientific comparative evaluation of the accuracy of truly non-invasive glucometers is not yet available. Thus, such a comparison can only be made based on data provided by research groups and manufacturers.

The feasibility of NIGM approaches is supported by a clinical test of another non-invasive system based on Raman spectroscopy reported recently²⁹. This NIGM system reached a MARD value of 19.2%. This compares well with the accuracy obtained with the system tested in this study, yet in a different study setup with a substantially longer calibration period of 26 days and home use without dedicated glucose swings.

Conclusions

We report here a clinical study on the evaluation of a truly non-invasive glucose sensor reaching an accuracy of MARD < 20% which is comparable to FDA-approved early stage CGMs. As a non-invasive measurement holds

many advantages compared to (minimally) invasive methods, the benefits for a device based on the demonstrated accuracy are sufficient for adjunctive use.

The measurements were based on the glucose absorption in the mid-infrared spectral region, at wavelengths between ~ 8 and $11\text{ }\mu\text{m}$. This “glucose fingerprint” (Fig. 1) is highly specific for the glucose molecule. Since the spectral features of glucose in this spectral range arise from fundamental C-C-O- and -O-H vibrations, their characteristics are more reliable than the overtones (harmonics) that are found in the near infrared between ~ 1 and $3\text{ }\mu\text{m}$, at the expense of more sophisticated light sources, optics and detection technology.

The external cavity-tuned QCLs used in the D-Base were originally developed as powerful and broadband tunable MID-IR sources for a wide range of scientific applications. A QCL optimized for the requirements of BG measurements could improve the performance as well as reduce the size and costs. Broad tunability over the entire glucose fingerprint (Fig. 1) is not an essential requirement but can be substituted by using selected wavelengths in the fingerprint spectral range. Using optimized arrays of QCL emitters that can be controlled individually with sufficient emission power could significantly reduce the measurement time to well below a minute and enable further miniaturization. Based on the gathered knowledge, DiaMonTech developed the first miniaturized dedicated glucose-QCL-array in cooperation with a German laser manufacturer. These arrays, which open the possibility for integration into wearable technology, could mitigate the drawbacks of the off-the-shelf QCLs used in the D-Base.

Based on these QCL arrays and learnings from the D-Base development, DiaMonTech has already miniaturized optics and electronics towards a handheld version “D-Pocket” as visualized in Fig. 8. The detection method in the D-Pocket is essentially the same as for D-Base. The D-Pocket has the footprint of a common smartphone. It is currently a prototype under development. The D-Pocket is planned as an individual companion for people with diabetes with the capacity for 30–50 measurements with one battery charge.

A small number of data points exhibit substantial deviations between the non-invasive measurement from the reference and are part of the investigations in the ongoing development work. Lab testing with the D-Base as well as further analysis of data from this study suggest that sources for signal instabilities for this non-invasive glucose measurement technique are the device itself, skin inhomogeneity, movement artefacts and artefacts from skin-care products. We expect to cope with the inhomogeneity of skin and influence of skin-care products by a more detailed analysis of the raw infrared spectra through modulation of the MIR pump laser. The aim is to identify sweat glands, scars or hair, substances on the surface of the skin and other potential sources for measurement inaccuracies. Reduction of the



Fig. 8 | Visualization of the handheld glucometer “D-Pocket”. With a form factor of a smartphone it will be portable and suitable for everyday life usage.

measuring time from a few min to the order of 10–20 s greatly reduces the risk of movement artefacts.

The data collected in this study will be used to further improve the algorithms with the aim of simplifying the calibration procedure and improving performance. As demonstrated, algorithm 2a, without recalibration on the session day, performed only slightly worse than the best algorithm, giving reason to believe that the algorithms can further be refined by reducing calibration requirements. With more data available, algorithms could be enhanced for a global calibration model for little to no (re-)calibration. We expect to collect this data with the handheld version “D-Pocket” that is currently finalized. The “D-Pocket” is planned as an individual companion for people with diabetes. As such, a trial with a larger number of devices used individually by the patients is planned. For the future, a wearable form factor based on a further miniaturized technology is in the concept phase.

In summary, we have successfully demonstrated the accuracy of a non-invasive glucose measurement technology. The study procedure was designed in a prospective fashion, i.e., measurements were done without prior knowledge of the reference values, aiming to a realistic situation of a diabetes patient. Based on MARD values, the Pearson correlation coefficient, the Bland and Altman analysis, and the consensus error grid analysis, this NIGM system exhibits the ability to follow changes of BG concentrations. The demonstrated performance opens new perspectives for better management of diabetes through realization of a reliable and accurate product for everyday use in the near future.

The device developed and used in this clinical validation (D-Base), and the handheld version under development (D-Pocket), do not require consumables such as disposable sensors, except for the test strips used for the calibration procedure with an invasive glucometer as described above. Yet, both allow an unlimited number of measurements to improve the glycemic condition of the diabetes patient. The issue of lifetime of these optical sensors (e.g., as compared to the maximum lifetime of CGM sensors of about 14 days) is not relevant. As for the lifetime of the entire sensor device,

we estimate based on lab tests a lifetime of >3 years limited only by standard wear and tear of an optical measurement device and a battery.

For the handheld version, preliminary cost estimates can be made based on a scalable production of the main component, the quantum cascade laser, the associated optics, and a lifetime of the device of several years. As an outcome, the costs over time for non-invasive glucose measurement using this technology will be strongly competitive to the costs for minimally invasive sensors.

Data availability

The complete source data for Figs. 5, 6 and 7 can be found in one Excel sheet in Supplementary Data. The complete report of the clinical test at the IfDT can be requested from the corresponding author.

Received: 25 October 2024; Accepted: 3 November 2025;

Published online: 15 November 2025

References

1. International Diabetes Federation. *IDF Diabetes Atlas: 11th Edition* <https://diabetesatlas.org> (International Diabetes Federation, 2025).
2. Petry, S. F., Petry, F. W., Petry, J. K., Gähth, S. & Heinemann, L. Diabetes technology and waste: a real-world study in a specialized practice in Germany. *J. Diab. Sci. Technol.* **2024**, 1–9 (2024).
3. Gonzales, W. V., Mobashsher, A. T. & Abbosh, A. The progress of glucose monitoring—a review of invasive to minimally and non-invasive techniques, devices and sensors. *Sensors* **19**, 800 (2019).
4. Pleus, S. et al. Documentation of skin-related issues associated with CGM use in the scientific literature. *Diabetes Technol. Ther.* **21**, <https://doi.org/10.1089/dia.2019.0171> (2019).
5. Di Fillippo, D., Sunstrum, F. N., Umar Khan, J. & Welsh, A. W. Non-invasive glucose sensing technologies and products: a comprehensive review for researchers and clinicians. *Sensors* **23**, 9130 (2023).
6. Hina, A. & Saadeh, W. Noninvasive blood glucose monitoring systems using near-infrared technology—a review. *Sensors* **22**, 4855 (2022).
7. Scholtes-Timmerman, M. J., Bijlsma, S., Fokkert, M. J., Slingerland, R. & van Veen, S. J. F. Raman spectroscopy as a promising tool for noninvasive point-of-care glucose monitoring. *J. Diab. Sci. Technol.* **8**, 974–979 (2014).
8. Pandey, R. et al. Noninvasive monitoring of blood glucose with Raman spectroscopy. *Acc. Chem. Res.* **2017** **50**, 264–272 (2017).
9. Lundsgaard-Nielsen S. M. et al. Critical-depth Raman spectroscopy enables home-use non-invasive glucose monitoring. *PLoS One* **13**, <https://doi.org/10.1371/journal.pone.0197134> (2018).
10. Pleitez, M. A., von Lilienfeld-Toal, H. & Mäntele, W. Infrared spectroscopic analysis of human interstitial fluid in vitro and in vivo using FT-IR spectroscopy and pulsed quantum cascade lasers (QCL): establishing a new approach to non-invasive glucose measurement. *Spectrochim. Acta Part A* **85**, 61–65 (2012).
11. Liakat, S. et al. Noninvasive in vivo glucose sensing on human subjects using mid-infrared light. *Biomed. Opt. Express* **5**, 2397–2404 (2014).
12. Kasahara, R., Kino, S., Soyama, S. & Matsuura, Y. Noninvasive glucose monitoring using mid-infrared absorption spectroscopy based on a few wavenumbers. *Biomed. Opt. Express* **9**, 289–302 (2017).
13. Uluc, N. et al. Non-invasive measurements of blood glucose levels by time-gating mid-infrared optoacoustic signals. *Nat. Metab.* **6**, 678–686 (2024).
14. Basu, A. et al. Time lag of glucose from intravascular to interstitial compartment in humans. *Diab.* **62**, 4083–4087 (2013).
15. Bauer A., et al. IR-spectroscopy of skin in vivo: optimal skin sites and properties for non-invasive glucose measurement by photoacoustic and photothermal spectroscopy. *J. Biophotonics* **11**. <https://doi.org/10.1002/jbio.201600261> (2018).

16. Pleitez, M. A. et al. Photothermal deflectometry enhanced by total internal reflection enables noninvasive glucose monitoring in human epidermis. *Analyst* **140**, 483–488 (2015).
17. Hertzberg, O., Bauer, A., Küderle, A., Pleitez, M. A. & Mäntele, W. Depth-selective photothermal IR spectroscopy of skin: potential application for non-invasive glucose measurement. *Analyst* **142**, 495–502 (2017).
18. Savitzky, A. & Golay, M. J. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**, 1627–1639 (1964).
19. Mujica, L. E., Rodellar, J., Fernandez, A. & Güemes, A. Q-statistic and T2-statistic PCA-based measures for damage assessment in structures. *Struct. Health Monit.* **10**, 539–553 (2010).
20. Lorber, A., Faber, K. & Kowalski, B. R. Net analyte signal calculation in multivariate calibration. *Anal. Chem.* **69**, 1620–1626 (1997).
21. Awad, M. & Khanna, R. Support vector regression. In: Efficient learning machines. Apress. https://doi.org/10.1007/978-1-4302-5990-9_4 (2015).
22. Bishop C. M. & Nasrabadi N. M. *Pattern Recognition and Machine Learning* (Springer, 2006).
23. Parkes, J. L., Slatin, S. L., Pardo, S. & Ginsberg, B. H. A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose". *Diab. Care.* **23**, 1143–1148 (2000).
24. Lubinski, T., Plotka, B., Janik, S., Canini, L. & Mäntele, W. Evaluation of a novel noninvasive blood glucose monitor based on mid-infrared quantum cascade laser technology and photothermal detection. *J. Diab. Sci. Technol.* **15**, 6–10 (2021).
25. Food and Drug Administration. Summary Of Safety And Effectiveness Data - Medtronic Guardian RT (PMA P980022/S011) Available at: https://www.accessdata.fda.gov/cdrh_docs/pdf/P980022S011B.pdf; Accessed date: 29 May 2024.
26. Food and Drug Administration. Summary Of Safety And Effectiveness Data - DexCom#STS# Continuous Glucose Monitoring System (PMA P050012) Available at: https://www.accessdata.fda.gov/cdrh_docs/pdf5/p050012b.pdf; Accessed date: 29 May 2024.
27. Freckmann, G. et al. Mean absolute relative difference of blood glucose monitoring systems and relationship to ISO 15197. *J. Diab. Sci. Technol.* **16**, 1089–1095 (2022).
28. Heinemann, L. et al. Benefits and limitations of MARD as a performance parameter for continuous glucose monitoring in the interstitial space. *J. Diab. Sci. Technol.* **14**, 135–150 (2020).
29. Pors, A. et al. Accurate post-calibration predictions for noninvasive glucose measurements in people using confocal Raman spectroscopy. *ACS Sens.* **8**, 1272–1279 (2023).

Acknowledgements

We thank Raja Bilal Ahmed, Philippa Clark, Jenny Kegel, Johannes Lutz, Yigit Mahsereci, Nikita Malyshev, Lars Melchior, Alexander Mittelstädt, Daniel Rojas (all at DiaMonTech) as well as /dev for preparing and

maintaining the D-Bases and making the study possible. We would also like to thank the IfDT for conducting this study and for the pleasant collaboration.

Author contributions

M.K., S.J., P.L. and M.S. prepared the non-invasive measurement technology for clinical validation and the general study design. L.C. and V.L. were responsible for the development of the signal processing and machine learning algorithms. T.L. supervised the study. W.M. wrote the first draft of the manuscript. All authors edited the manuscript.

Competing interests

M.K. is CTO, S.J. is COO, T.L. is CEO, and W.M. is CSO of DiaMonTech AG. T.L. and W.M. are co-founders of DiaMonTech AG. L.C., P.L., V.L. and M.S. are employees of DiaMonTech.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43856-025-01241-7>.

Correspondence and requests for materials should be addressed to Werner Mäntele.

Peer review information *Communications Medicine* thanks Anas El Fathi, Pradyut Kumar Sanki and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025