Maximilian Lübke*, Angelika S. Thalmayer, Vincent Lassen, Norman Franchi, and Bernhard Etzlinger

The Demand and the Path to Trustworthy Continuous Blood Glucose Monitoring

https://doi.org/10.1515/cdbme-2025-0222

Abstract:
With the growing role of telemedicine and artificial intelligence (AI) in healthcare, trust and trustworthiness have become essential considerations — especially for systems relying on medical sensors and patient-specific data. Using Continuous Glucose Monitoring (CGM) as a representative case, this work explores how technical metrics (e.g., accuracy, security, robustness) and subjective user perceptions can be combined to evaluate trustworthiness in the context of Diabetes care, considering different stakeholders' perspectives. We highlight key risks across the CGM lifecycle and propose a trustworthiness factor that enables dynamic, bi-directional interaction between system assessments and user trust. This approach offers a foundation for making trustworthiness tangible and adaptable in future AI-driven medical applications.

Keywords: CGM, trust, trustworthiness, wearables.

1 Motivation

Diabetes mellitus is increasing worldwide, with millions affected and rising numbers, especially in industrialised nations. In parallel, research has accelerated, and commercial products such as continuous glucose monitoring (CGM) systems have emerged. These systems, now in their fourth generation and beyond, offer significant relief in therapy. However, current diabetes care remains far from optimal. Key stakeholders expect substantial digitalisation in diabetes care by 2030 [1]. This includes telemedicine using artificial intelligence (AI) or avatars acting as doctors to enhance routine monitoring, emergency responses, and therapy adoption. While gains in therapy quality, reduced patient burden, and lower healthcare costs are strong motivators, trust in these complex systems remains a major concern. A critical question arises: to what extent can stakeholders in diabetes care trust a telemedicine application?

***Corresponding author: Maximilian Lübke,**
Friedrich-Alexander-Universität Erlangen-Nürnberg, Institute of Smart Electronics and Systems, 91058 Erlangen, Germany,
E-mail: maximilian.luebke@fau.de
Angelika S. Thalmayer, Vincent Lassen, Norman Franchi,
Friedrich-Alexander-Universität Erlangen-Nürnberg, Institute of Smart Electronics and Systems, 91058 Erlangen, Germany
Bernhard Etzlinger, Johannes Kepler University Linz, Institute for Communications Engineering and RF-Systems, 4040 Linz, Austria

Obviously, as trust is the subjective feeling of a human, a technical system can only provide a evidence-based indication of trust, referred to as *trustworthiness* [2]. In this article, we approach the question above through the quantification of trustworthiness in state-of-the-art systems used in Type 1 Diabetes mellitus therapy. In such systems, a distinction is made between closed-loop systems, which are based on the patient's diagnosis and lead to direct therapy (e.g. via insulin pumps), and the classic therapy of manual insulin injections. Hereby, we focus primarily on the latter one. As the insulin therapy is heavily dependent on the current blood glucose level (BGL), which in turn is measured by the CGM system, these measuring systems must be classified as particularly critical. If too much insulin is injected, there is a risk of hypoglycemia, which in the worst case can even lead to the patient's death. If the BGL is too high over a long period of time, there is a risk of most likely long-term effects, however in extreme cases can also lead to coma, but usually leads to serious secondary diseases (coronary system, feet, kidneys).

The results of the diagnostics are, therefore, directly decisive for the patient's everyday life, and in the interest of all stakeholders – that are persons with diabetes, healthcare professionals, manufacturers, and insurances. In the following, we examine trust and trustworthiness of CGM sensors. We include various scenarios and influencing factors (sensor, signal processing, and communication domains) that already exist in today's commercial systems and expand our analyses with regard to future trends (e.g. ever-increase in usage of AI, telemedicine) [3, 4]. Trustworthy environments in the context of telemedicine is hereby also reported as one key use case family in the 6G context [5]. These environments need technological ecosystems that prioritize security, privacy, and reliability, especially for human-centric services, where the safety of human life is paramount. Our analysis gives the reader an initial insight into the concept of combining trust and trustworthiness for medical - CGM - sensors.

2 Scenarios and Definitions

To design and operate trustworthy systems, a clear understanding of trustworthiness must first be established. This understanding can vary significantly between use cases, as stakeholder needs differ in priority. We therefore outline the key considerations and definitions below. Access to trustwor-

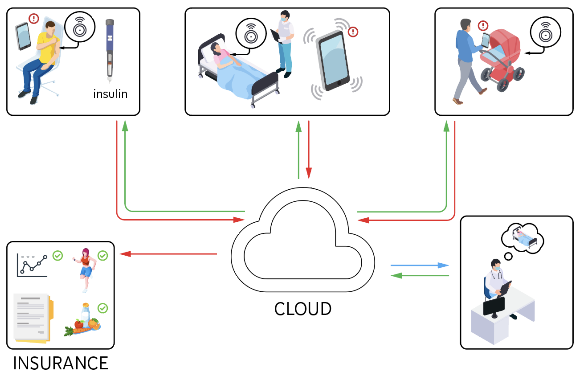

Fig. 1: CGM scenarios, including the perspective of different stakeholders (patient, carer, doctor and insurances).

thy sensor system evaluations and therapy recommendations should be readily available to patients, caregivers, and doctors. This is particularly valuable in situations where direct feedback from patients is limited, such as when parents are caring for infants or when hospital staff without specialized diabetic training are looking after patients. Especially during and after surgical procedures, when aftercare may be provided by staff unfamiliar with diabetes management, trustworthy diagnostic and treatment recommendations could significantly reduce stress for both, patients and caregivers.

2.1 Considered CGM Scenarios

Stakeholders directly involved in insulin treatment include the patient and healthcare or nursing staff (top row in Figure 1), all relying on system outputs processed and shown on the coupled handhelds. Additionally, there are several stakeholders connected via the cloud, as shown in Figure 1. Physicians depend on CGM data to adjust therapy and prevent emergencies, while insurers monitor patient compliance via system outputs. From this interconnected CGM system, we focus on the "first metre"—comprising the sensor itself, signal processing, and wireless communication between the sensor and the receiving device (e. g., a mobile phone).

2.2 Trustworthiness Definition

While early concepts of trustworthiness analyses trace back to military software in the 80's, it was recently recognized as an essential concept in interdependent systems. Not surprisingly, it was hence raised in the IoT and cyber-physical systems community [6] and then further prominently transferred to AI and cellular communication systems [7]. Thereby, the following definition is widely adopted (with minor variations) [6]: "Trustworthiness is the demonstrable likelihood that the system performs according to designed behavior under any set of conditions as evidenced by characteristics including, but not limited to, safety, security, privacy, reliability and resilience."

While each characteristic can be measured through established metrics, a meaningful adaption to the use-case is required [8]. A resulting trustworthiness value will provide all involved parties of Diabetes care an objective parameter on how well the patient/carer or doctor can rely on system outputs. In the following we outline its application for CGM open-loop systems.

2.3 Trustworthiness in CGM Systems

In the scenarios mentioned above, faulty, inconsistent, or slightly varying data or glucose predictions can significantly impact patient health. Beyond reliability, security and privacy, additional concerns such as usability and robustness arise. Currently, stakeholders often rely blindly on system outputs, leaving them alone with their subjective trust feeling when interpreting the system output. A computed trustworthiness value, based on measurable evidence during system operation, can indicate system vulnerabilities or failures (e.g., a sensor no longer adhering properly to the skin) and enable proactive responses in critical situations (e.g., issuing warnings near operational limits). To successfully implement trustworthiness monitoring, all stakeholders are required to include their definitions at all functional levels of a properly operating CGM system. In this article, we focus on the level of data extraction.

2.4 Trust and Trustworthiness

As outlined, trust is the subjective feeling of a human, while trustworthiness refers to the technical measure. On the one hand, the smartest technical system will only be able to operate in foreseen situations. On the other hand, there is nothing more individual than the patients metabolism and the human perception. Thus, the trustworthiness assessment cannot be treated as a static procedure applied once to a given system. The assessment has to be constantly adapted to the user through an interface that enables to dynamically adjust threshold levels in the trustworthiness evaluation. Establishing this interface requires interdisciplinarity and remains out of the scope of this article.

3 Towards CGM Metrics

A key challenge in designing trustworthiness assessments lies in the selection of appropriate metrics. A systematic approach proposed in [9] recommends choosing metrics based on the system's specific threats and is applied in the following.

3.1 Sensor Domain

Accuracy: The accuracy of the measurement is certainly the core aspect of a reliable diagnosis that leads to a reliable therapy. In the case of CGM, however, we are dealing with a very individual patient metabolism, which depends on the daily form and hormone levels and many other factors. Thus, traditional accuracy metrics of sensor readings –while still bearing high importance for product safety– have limited meaning in the context of trustworthiness assessment during operation. On the other hand, these traditional metrics form the basis for the FDA in the USA and the EMA in the EU to assess the fulfillment of accuracy requirements for blood glucose meters. The Clarke error grid [10, Fig. 14] illustrates mismatches between actual BGLs and the measured sensor reading depending on the patients metabolic pattern. The error range is partitioned into 5 zones, labeled with capital letters from A to E. Clinically valid measurements fall within zone A of the error grid, where the tolerance is below 20 %, or both values are below 70 mg/dL. Other zones (B-E) in the grid range from clinically uncritical to extremely dangerous, depending on the discrepancy between the values: This is particularly critical for zone D, resulting in dangerous discrepancies that could fail to detect and treat true blood glucose conditions, and zone E with extremely dangerous mistakes where the blood glucose level will be incorrectly treated. Translating the Clarke error grid to the CGM sensor in operation, we identify the two critical accuracy states to be assessed: Complete sensor malfunction, and zones of hypo- and hyperglycemia.

Resilience: The potential for a sensor app to crash, or errors in the BGL measurement on the sensor side require the sensor to be recalibrated. Resilience measures how quickly one can reach an accurate BGL measurement (again) and how swiftly a communication link is established. Together, these aspects influence the overall effectiveness and user satisfaction. Effective metrics are mean time to recover and mean time between failures, that are easily measured through the (internally or externally triggered) recalibration process.

Robustness and Usability: Both are critical aspects when relying on the technical system. There are several important factors to monitor. Technological incapability with regard to software updates or software/hardware errors resulting in high energy consumption by the handheld device and improper software deployment, can significantly impact usability. Additionally, the aesthetic appeal of the sensor (patients prefer not to be identified as such) and the ease of access to the data and/or the processed metrics, like estimated HbA1c value, for patients, carers and doctors are also crucial factors. We expect that insurances are also highly interested in this aspect.

3.2 Signal Processing Domain

The signal processing of the sensors takes place on the sensors themselves, on the handhelds, and in the cloud and have impact depending on the target horizon. Beside determining the current BGL, a relatively powerful feature of current CGM sensors is to provide a trend analysis and forecast. It indicates whether the BGL is likely to rise (steeply, slowly), remain constant or fall (also steeply or slowly). Thus, this tool makes it possible to take early countermeasures, i.e. sugar intake to prevent hypoglycemia or insulin injection to counteract the consequences of hyperglycemia. The underlying methods to get to such a trend analysis are under continuous development with an ever-increasing trend towards the use of AI. We address some aspects of AI in the medical/CGM context below.

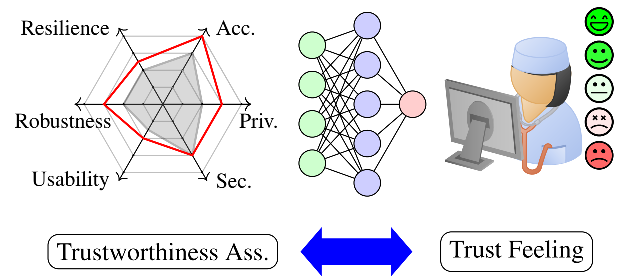

Fig. 2: Proposed bi-directional interface to achieve a trustworthy and trusted CGM analysis, respectively.

Trustworthy AI: In 2020, the High-Level Expert Group on Artificial Intelligence of the European Commission published the Assessment List for Trustworthy AI, which served as a foundational document for the AI Act—a legally binding regulation aimed at ensuring the safety and protection of fundamental rights in AI systems across the EU market. While the regulatory framework does not yet provide detailed criteria for evaluating AI trustworthiness, EU member states are actively developing certification procedures. Notably, *Trusted AI* by TÜV Austria is the world's first machine learning certification scheme, offering a binary classification of algorithms as either trustworthy or not. Similar initiatives are expected to follow in other countries. However, dynamic or runtime trustworthiness assessment of AI algorithms is not currently addressed. Consequently, the overall certification process for CGM systems prior to market approval will need to include all embedded AI components, formally verifying them as trustworthy AI.

Subjective Trend Misalignment: Trend analysis is essential for optimizing BGL management and guiding treatment recommendations. However, inaccurate trends combined with faulty BGL measurements can lead to incorrect interventions, such as excessive insulin doses. This poses a danger if e.g. the actual BGL levels are normal or low, and misinterpreted by patients or caregivers. Patients often intuit their BGL changes, and discrepancies between sensor readings and patients' perceptions can erode trust.

3.3 Communication System Domain

There are two communication paths: on one hand, the transmission between the sensor and handheld device, and on the other hand, the transfer to a cloud. The communication from the sensor to the smartphone is through bluetooth (i.e., the CGM service specification), while the cloud connection through cellular communication and internet connection. Both communications entail risks, which we will summarize below:

Privacy and Security: Privacy and security concerns in handling data from sensors to databases and eventually to cloud storage can be significant. On the privacy side, there is the potential for misuse of data, such as using data to identify individuals or creating lifestyle profiles that infringe on personal privacy. On the security side, there is the risk of incorrect data being transmitted to the handheld device or stored in the database, which could lead to improper therapy decisions. These risks necessitate a thorough analysis to ensure both privacy and security are adequately addressed.

4 User Informed Evaluation

Metrics related to the characteristics highlighted above (i.e., accuracy, resilience, robustness, usability, security, and privacy) typically yield values that are meaningful primarily to technical experts. To make these metrics interpretable for non-expert users of CGM systems, they must be aggregated into compound values and compared against predefined thresholds. This process of data interpretation reflects how thoroughly CGM manufacturers have anticipated potential scenarios throughout the system's lifecycle and embedded appropriate online evaluation mechanisms during the design phase. However, to also address unforeseen situations, a feedback loop from the user to the CGM system is essential. If a user perceives a significant mismatch between their trust perception and the system's trustworthiness assessment, adaptations—such as adjustments to threshold levels—can be triggered, as illustrated in Figure 2. To sustain user trust throughout the system's lifecycle, the interface between real-time trustworthiness assessment and the user must therefore be bi-directional, enabling both interpretation and influence.

5 Conclusion

Trustworthiness is a critical requirement across the entire telemedicine pipeline, especially in AI-supported diagnosis and therapy recommendation systems. From the initial collection of physiological data by sensors, through transmission and storage, to analysis and therapeutic output, each step must be secure, reliable, and interpretable to maintain system integrity and user confidence. Loss of trust may result from a variety of factors, including data inconsistencies, sensor displacement, poor usability, or concerns about security, and safety.

In this work, we have identified and analyzed potential risks to trustworthiness within continuous glucose monitoring (CGM) systems, and leveraged key aspects for online evaluations by means of metrics. While our focus has been on CGM, the outlined principles are broadly applicable to other telemedical applications, particularly those involving handheld or wearable technologies. As an initial step toward operationalization trust, we have proposed a trustworthiness factor that integrates both objective system metrics and subjective user perceptions. Future work will aim to refine this approach, particularly in the context of increasingly AI-controlled therapies, where adaptive, user-centred trust mechanisms will be vital.

Acknowledgment: This work was, in part, supported by the Federal Ministry of Education and Research, Germany (BMBF) as part of the project "6G Platform Germany" under the grant number 16KISK048. It is also partly based upon work from COST Action 6G-PHYSEC (CA22168), supported by COST (European Cooperation in Science and Technology). The work contributes to the research within the 6G-Valley innovation cluster. The authors acknowledge the use of OpenAI's ChatGPT to assist with grammar and language refinement in preparing this manuscript.

References

[1] Nolan J J, Valentin T. A Vision for Digitally Enabled Diabetes Care in Europe. Views of Leading Stakeholders. TechMed Europe, White Paper; 2022.
[2] Cho J-H, Xu S, Hurley P M, Mackay M, Benjamin T, Beaumont M. STRAM: Measuring the trustworthiness of computer-based systems. ACM New York, NY, USA; 2019.
[3] Aruna Devi B, Karthik N. Machine Learning-Based Trustworthy Diabetes Prediction Model. 2024 16th International Conference on COMmunication Systems & NETworkS (COMSNETS) 2024; Bengaluru, India; 400-402.
[4] Savchuk D, Doroshenko A. Chapter 3 - Explainable AI methods to increase trustworthiness in healthcare. Responsible and Explainable Artificial Intelligence in Healthcare 2025; Academic Press; 55-89.
[5] Franchi N, Dressler F, et al. German Perspective on 6G – Use Cases, Technical Building Blocks and Requirements. Insights by the 6G Platform Germany, White Paper. FAU Studien aus der Elektrotechnik Band 28.Erlangen.
[6] Cyber-Physical Systems Public Working Group (CPS PWG). Framework for Cyber-Physical Systems. National Institute of Standards and Technology (NIST); 2016.
[7] Fettweis G P, Boche H. On 6G and trustworthiness. Communications of the ACM, 2022.
[8] Osiorio D, Etzlinger B, Karamachoski J. Trustworthy 6G: Misconceptions, Attributes, and the Labeling Approach. Authorea Preprints, 2024.
[9] Peterseil P, Etzlinger B, Horáček J, Khanzadeh R, Springer A. Trustworthiness for an Ultra-Wideband Localization Service. Sensors, 2024.
[10] Xue Y, Thalmayer AS, Zeising S, Fischer G, Lübke M. Commercial and Scientific Solutions for Blood Glucose Monitoring—A Review. Sensors 2022; 22:425.