Patient Related Outcome Measures

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

# Towards Standardization of CGM Performance Studies: The QUSS-CGM Questionnaire for Assessing User Satisfaction

Anne Beltzer[1], Marta Gil Miró[1], Manuel Eichenlaub[1], Delia Waldenmaier[1], Cornelia Haug[1], Dominic Ehrmann[2,3], Guido Freckmann[1]

[1]Institute for Diabetes Technology Ulm, Ulm, Germany; [2]Research Institute Diabetes Academy Mergentheim (FIDAM), Bad Mergentheim, Germany; [3]Department of Clinical Psychology and Psychotherapy, University of Bamberg, Bamberg, Germany

Correspondence: Anne Beltzer, Institute for Diabetes Technology Ulm, Lise-Meitner-Street 8/2, Ulm, 89081, Germany, Email anne.beltzer@idt-ulm.de



**Purpose:** User satisfaction and ease of use of continuous glucose monitoring (CGM) systems are key factors in patients' device acceptance. CGM user satisfaction is often assessed through questionnaires, but item selection varies widely across studies. The aim of this study was to design, develop and validate a Questionnaire for User Satisfaction Standardized for CGM performance studies (QUSS-CGM).

**Methods:** Selection of attributes and design of questionnaire items was based on a systematic literature search of publications on CGM performance evaluation studies. Content and response process validation of a draft-questionnaire was performed by experts (n=9) and people with diabetes (n=10), respectively. The resulting German pre-QUSS-CGM questionnaire underwent validation in two CGM performance studies ("pilot" studies) performed between June and August of 2024, via a pooled psychometric evaluation (exploratory factor analysis (EFA) and reliability) of n=126 questionnaires from these studies, followed by bidirectional translation to English.

**Results:** Two hundred and five items on user satisfaction in CGM performance studies were identified by systematic literature search and classified into six attributes according to their content. Items were summarized in a 25-item draft-questionnaire on a 5-point Likert scale. Content and face validity were considered acceptable with a scale-level content validity index (S-CVI/Ave) of 0.90 and a scale-level face validity index (S-FVI/Ave) of 0.93, both based on the average method. EFA revealed a two-factor structure for the final QUSS-CGM questionnaire summarized to 11 items, demonstrating high internal consistency (Cronbach's α of 0.84).

**Conclusion:** The QUSS-CGM was designed, developed, and validated as a reliable and standardized tool to measure user satisfaction in CGM performance evaluation studies.

**Keywords:** user satisfaction, questionnaire, content validation, continuous glucose monitoring, explorative factor analysis, performance

## Introduction

Continuous glucose monitoring (CGM) systems are part of everyday life for many people with diabetes and their increased accuracy in recent years allows them to be used for therapy decision-making.[1] Therefore, adequate device performance is indispensable to ensure reliable diabetes management outcomes. However, studies to determine clinical performance of a CGM system vary tremendously in terms of study design, selection of comparator method, experimental procedures, or statistical analysis, making it difficult both to interpret the data (eg, accuracy) and to compare devices among different studies.[2] Efforts are being made under the direction of the International Federation of Clinical Chemistry (IFCC) Working Group on CGM to standardize the reporting and procedures for assessing the performance of CGM systems.[2–4]

Apart from accuracy and reliability, user satisfaction remains an integral component of CGM performance assessment and constitutes a critical determinant in patient adherence to long-term device utilization. High levels of patient satisfaction have been linked to more frequent CGM use and are key to patient adherence.[5] A widely used tool to assess user satisfaction are questionnaires,[6] as shown by the literature review of CGM performance studies from Freckmann



Received: 21 July 2025
Accepted: 9 December 2025
Published: 19 December 2025

et al.[2] Several validated questionnaires for glucose monitoring user satisfaction already exist, eg, the Glucose Monitoring Experiences Questionnaire (GME-Q),[7] the Glucose Monitoring System Satisfaction Survey (GMSS),[8] the CGM Satisfaction Scale (CGM-SAT) and the Glucose Monitoring Survey (GMS).[9] These questionnaires cover general aspects of usability and user satisfaction with a major focus on quality of life and patient-reported outcomes during long-term usage of a CGM system, as well as its therapeutic impact.[8,9] However, in CGM performance evaluation studies, the wearing time of these minimally invasive systems is comparatively short (generally no longer than two weeks) and factors such as sensor handling, pain or general comfort are primarily addressed rather than long-term experience. In addition, the therapeutic impact cannot be judged by the participant of CGM performance studies since they are usually only allowed to use their own system for therapy decision-making. When reviewing other questionnaires used in CGM performance studies, the selection of questions and surveyed aspects was found to be unsystematic. To our knowledge, there is no validated method specifically tailored to assess user satisfaction with minimally invasive CGM systems in the context of performance studies.

The development of a standardized user satisfaction questionnaire requires a systematic validation approach to ensure high-quality results. Therefore, the objective of this work was to design, develop, and validate the Questionnaire for User Satisfaction Standardized for CGM performance studies (QUSS-CGM). This involved compiling relevant satisfaction questions referred to as items through a systematic literature review, identifying key topics known as attributes, and conducting a structured validation process, including content validation with experts and response process validation with individuals with diabetes to assess item relevance and comprehensibility. A psychometric evaluation followed, incorporating item selection based on corrected item total correlation, exploratory factor analysis (EFA) and reliability assessment. Finally, the QUSS-CGM was translated to English using a forward-backward translation method.

# Materials and Methods

## Context

The intended target population of the questionnaire are participants (people with type 1 and 2 diabetes aged ≥ 18 years) of two minimally invasive CGM performance evaluation studies.[10,11] Furthermore, the number of questions should be chosen such that the questionnaire can be completed in five to ten minutes. A schematic depiction of the QUSS-CGM development process is shown in Figure 1.

## Questionnaire Design and Item Development

The defined attributes were based on a sub-analysis of a systematic literature search conducted in the MEDLINE database by Freckmann et al,[2] which included 129 CGM performance studies published between 2002 and 2022. Key inclusion criteria were that studies evaluated ISF-based CGM systems in adults with type 1 or type 2 diabetes under routine care conditions. To retrieve more recent scientific publications on this topic, the search was extended by the time period from 01-Jan-2023 to 28-Aug-2023. All publications meeting the criteria defined by Freckmann et al[2] and containing relevant content on usability and user satisfaction were included in the analysis to gather items and define the categorizing attributes.

Identified items were grouped by content and assigned to different attributes. Covering the content of each attribute, concise items were positively and negatively phrased and organized in a two-part draft-questionnaire with items for the beginning and end of the study. A 5-point Likert scale was chosen as the response option for the questionnaire (1- “strongly disagree”; 2- “disagree”; 3- “partly agree”; 4-“agree”; 5- “strongly agree”), as it was often used in the questionnaires identified through the systematic search.

## Item Selection and Validation

Content validation was performed to confirm that the items mirrored the intended focus and were relevant to the objective of the questionnaire. A panel of experts (n=9) in the field of CGM use and performance evaluation, selected based on their publication track record and experience as investigators or physicians in CGM studies, were instructed to rate the defined draft-questionnaire items according to their relevance and clarity, separately, in addition to suggesting

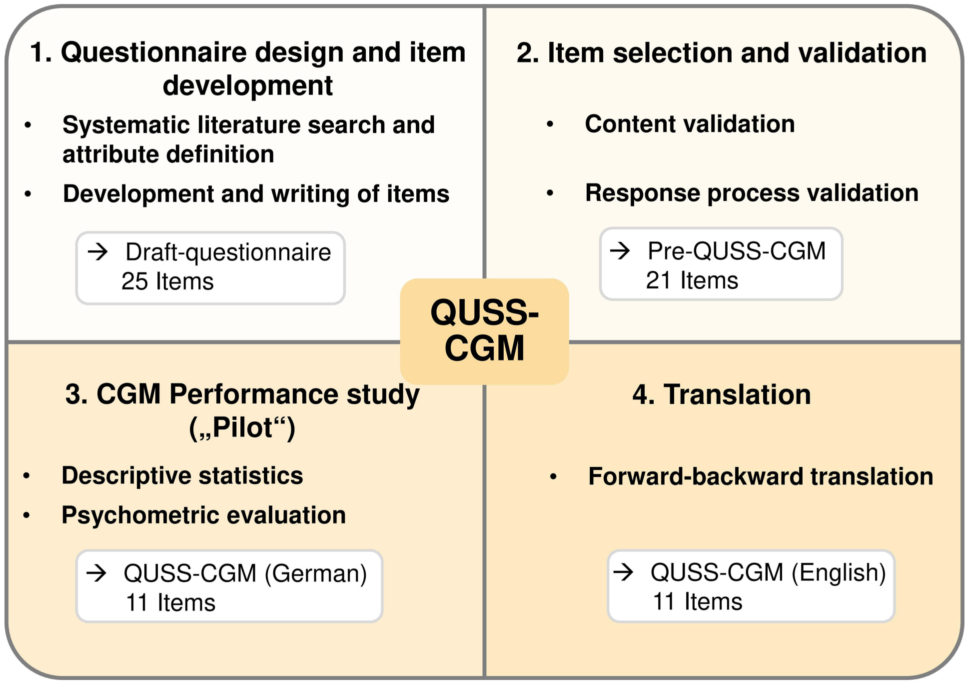


**Figure 1** Main steps followed to design, develop and validate the user satisfaction questionnaire. Questionnaire for user satisfaction standardized for CGM (QUSS-CGM).

phrasing changes.[12] The established scoring system used consisted of the item-level content validity index (I-CVI) and the overall, scale-level content validity index based on the average method (S-CVI/Ave).[12] Items with an I-CVI <0.78 were excluded,[12] and the remaining items were rephrased according to experts' suggestions. An overall S-CVI/Ave score of >0.80 was considered to be acceptable.[13]

Response process validation was conducted by a group of people with diabetes (n=10), reflecting the intended target of the questionnaire, to determine the face validity of the items that passed the content validation, evading potential ambiguities.[14] The comprehensibility and clarity of the items were rated, and suggestions could be made for clearer wording.[14] The established the item-level face validity index (I-FVI) and the overall, scale-level face validity index based on the average method (S-FVI/Ave) scoring system was used.[14] Items with an I-FVI <0.80 were revised and rephrased according to the respondents' suggestions.[14] An overall S-FVI/Ave score of >0.83 was considered to be acceptable.[15]

## Pilot Study and Psychometric Evaluation

Data from two CGM performance studies[10,11] performed between June and August of 2024 referred to as "Pilot studies" were pooled to increase the sample size and strengthen the statistical significance. For both studies, ethical approval was obtained before the start and they comprised a total of 54 participants, each of whom wore 2 or 3 distinct CGM systems concurrently. All participants provided written informed consent prior to enrollment. Four different systems were tested, one of them in both studies. Subjects completed one questionnaire per system, yielding 132 questionnaires. Questionnaires not fully answered (n=6) were excluded from the analysis. Floor and ceiling effects were investigated (when >15% of the answers focus on the worst and best ratings available on the scale, respectively) to assess the sensitivity of the questionnaire to detect differences at the lower and upper ends of the scale.

Item selection for the EFA was conducted through an iterative process, based on their corrected item total correlations, those below the cut-off of 0.2 were deleted. Items were selected until there was no substantial increase in Cronbach's α without loss of information. After assessing the adequacy of the sample with the Kaiser–Meyer–Olkin (KMO, >0.5) measure and the Bartlett sphericity test,[16] EFA was carried out to determine the factorial structure of the questionnaire.[17] The number of factors defining the structure of the questionnaire was established using the Kaiser criterion (eigenvalues >1), in addition to an inspection of the Scree Plot using the elbow criterion.[16] Factors were extracted through principal axis factoring using oblique rotation (Oblimin) to simplify the structure. Items were classified under each factor according to their corresponding factor loading score (>|0.4|), cross-loadings for more than one factor were assessed, as well as items with insufficient factor loading scores. Afterwards, questionnaire's internal reliability was evaluated with the Cronbach's α coefficient, with values above 0.7 considered as good.[17] Psychometric evaluation consisting of EFA and reliability assessment was performed using SPSS software version 29.

To interpret the outcomes of the questionnaire, a scoring standardization yielding one score per CGM system was carried out. To compute the total score, the established 5-point Likert scale responses (1–5) were recoded to a 0–4 range. Reverse scoring was applied to negatively worded items and similarly converted to the 0–4 range to maintain scoring consistency. The total score was linearly transformed to a scale from 0–100, with higher scores reflecting greater user satisfaction. A scale based on the percentile calculation was defined.[17]

## Translation

The development and validation of the QUSS-CGM was done in German since this was the mother tongue of the participants in the CGM performance studies. Hence, a validated English translation was developed using a bidirectional translation protocol, where native German bilingual experts translated the German version to English and native English bilingual experts translated the items back to German. Afterwards, the versions were reconciled to avoid cross-cultural differences and establish the final translation.

# Results

## Systematic Literature Search

The systematic literature search on CGM performance studies yielded a total of 130 articles (n=129 from 2002 to 2022,[2] n=1 from extended search in 2023), of which 31 manuscripts[18–48] contained relevant content in the form of user satisfaction questionnaires or individual items. Four of these articles were excluded since one evaluated a combined device for insulin infusion and glucose sensing[27] and three focused on an implantable CGM system.[31,38,39] Twenty-seven publications containing a total of 205 items were included (Figure 2a). Items were classified according to their content; thereafter, six attributes were defined: sensor handling, ease of use, digital interface, experience, comfort and reliability. Defined attributes were almost evenly represented in the publications, with "sensor handling" being the most frequent attribute and digital interface the least frequent (Figure 2b). This classification was based on the unmodified items, meaning that duplicates and equivalent items were all included.

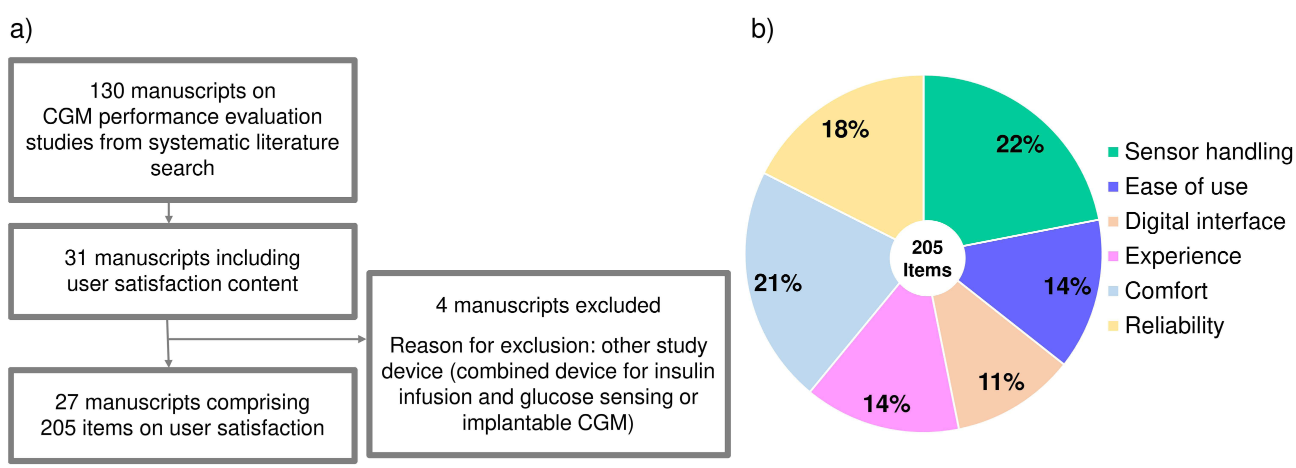


**Figure 2** (**a**) Literature selection process, based on the systematic literature search published by Freckmann et al 2023;[2] (**b**) Classification of the 205 raw items to six attributes compiled from the user satisfaction surveys used in the 27 CGM performance studies screened.

## Item Development

The 205 items pooled from the systematic literature search were classified into six defined attributes (sensor handling, ease of use, digital interface, experience, comfort, and reliability). After removing duplicates and consolidating similar items, 25 unique items remained, forming a draft questionnaire that exemplified the content collected through the literature research, following the methodology described by Yusoff et al (2021)[17] (Figure 3). This substantial reduction from 205 to 25 items was primarily due to the high number of duplicates and repetitive items across publications. Items 1 to 5 were designed to be asked after sensor insertion and items 6 to 25 at the end after sensor removal.

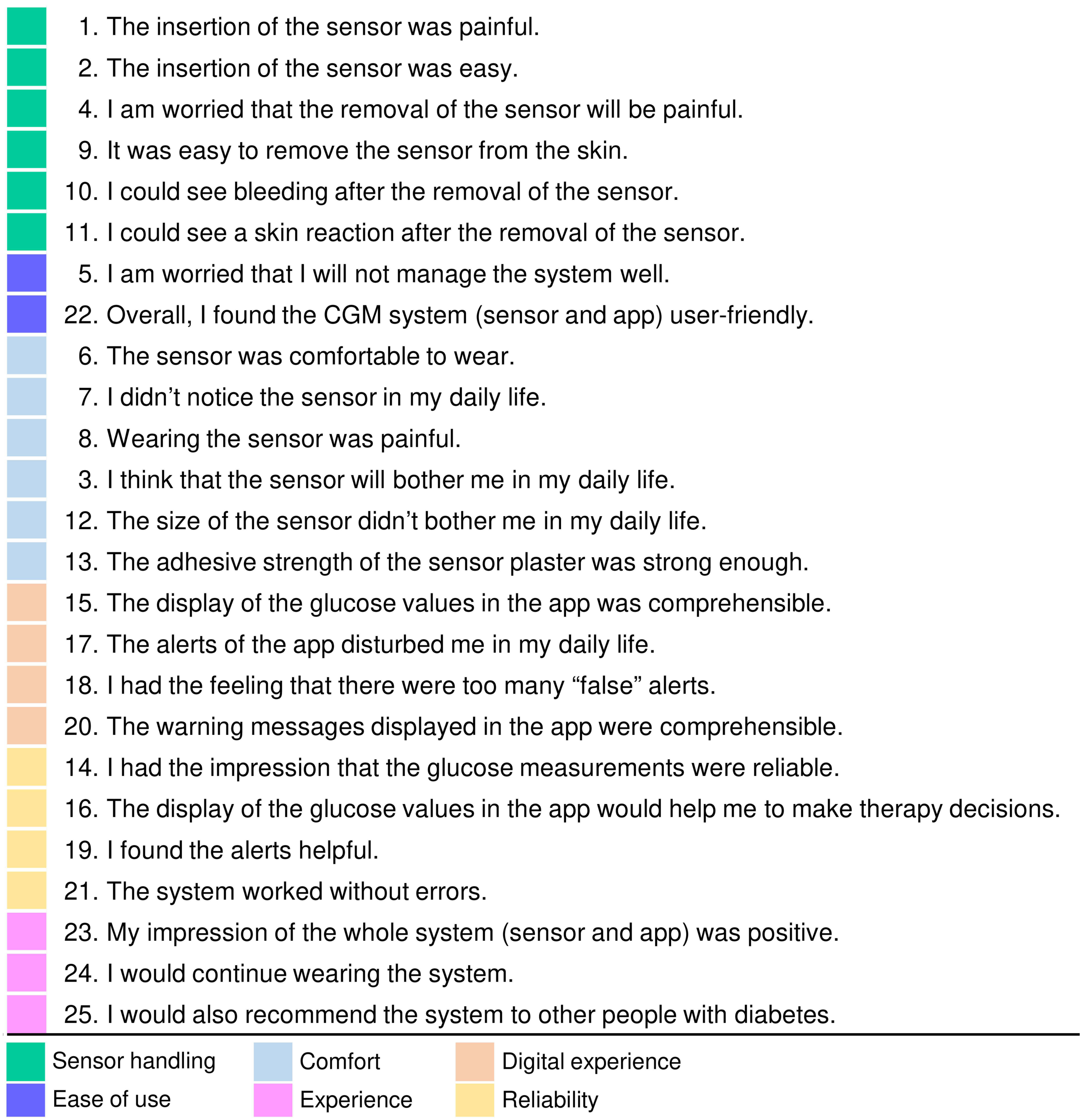


**Figure 3** Items constituting the draft-questionnaire before the validation steps; the 25 items were classified according to the attributes defined from the literature search (represented in different colors). Each item has a number assigned, corresponding to the order in which the items are planned to be answered.

The results of the content validation process of the 25 items by nine experts from the diabetes technology field are shown in Table 1. Twenty-one of the 25 initial items were scored with an I-CVI equal to or higher than 0.78; nonetheless, four items (4, 10, 16 and 17) did not reach the cut-off score and were excluded (Table 1). Ratings on clarity and rephrasing suggestions were taken into consideration, and minor item modifications were applied accordingly. The overall S-CVI/Ave of the remaining 21 items was 0.90, indicating an acceptable content validity.

Table 1 also shows the results of the face validity process of the remaining 21 items performed by ten people with diabetes. Nineteen of the total 21 items were scored over 0.80, whereas three received a score equal to or lower than 0.80, indicating that there could be ambiguities. These items were marginally modified according to the suggestions made by the respondents. The overall S-FVI/Ave was 0.93, indicating an acceptable face validity.

## Participant Characteristics of Pilot Studies

Among the 54 participants who completed the questionnaire, 65% were male, with a mean age of 51 years (min: 22, max: 71). Participants mostly had type 1 diabetes (96%) with a mean disease duration of 27 years (min: 1, max: 58). Moreover, the majority of participants (96%) reported prior experience with CGM systems.

**Table 1** Calculation of Item-Level Content Validity Indexes (I-CVIs) and Scale-Level Content Validity Index Based on the Average Method (S-CVI/Ave) of the 25 Items, Done by 9 Experts, Followed by the Calculation of Item-Level Face Validity Indexes (I-FVs) and Scale-Level Face Validity Index Based on the Average Method (S-FVI/Ave) of the 21 Items Remaining After Content Validity, Done by 10 Subjects; the S-CVI/Ave Is 0.90 and the S-FVI/Ave Is 0.93. *I-CVI Cut-off Value Is 0.78 and Therefore Not Met by Items 4, 10, 16 and 17. ¥items with an I-FVI Score Lower Than 0.80 Were Revised and Rephrased

| Content Validity | | | Face Validity | | |
|---|---|---|---|---|---|
| **Items** | **Experts in Agreement** | **I-CVIs** | **Items** | **Raters in Agreement** | **I-FVIs** |
| 1 | 9 | 1.00 | 1 | 10 | 1.00 |
| 2 | 9 | 1.00 | 2 | 10 | 1.00 |
| 3 | 8 | 0.89 | 3 | 9 | 0.90 |
| 4 | 5 | 0.56* | | | |
| 5 | 7 | 0.78 | 4 | 9 | 0.90 |
| 6 | 9 | 1.00 | 5 | 10 | 1.00 |
| 7 | 8 | 0.89 | 6 | 8 | 0.80¥ |
| 8 | 7 | 0.78 | 7 | 10 | 1.00 |
| 9 | 8 | 0.89 | 8 | 10 | 1.00 |
| 10 | 4 | 0.44* | | | |
| 11 | 7 | 0.78 | 9 | 8 | 0.80¥ |
| 12 | 8 | 0.89 | 10 | 9 | 0.90 |
| 13 | 9 | 1.00 | 11 | 9 | 0.90 |
| 14 | 7 | 0.78 | 12 | 9 | 0.90 |
| 15 | 9 | 1.00 | 13 | 9 | 0.90 |
| 16 | 5 | 0.56* | | | |
| 17 | 6 | 0.67* | | | |
| 18 | 7 | 0.78 | 14 | 9 | 0.90 |
| 19 | 8 | 0.89 | 15 | 10 | 1.00 |
| 20 | 9 | 1.00 | 16 | 10 | 1.00 |
| 21 | 8 | 0.89 | 17 | 7 | 0.70¥ |
| 22 | 9 | 1.00 | 18 | 10 | 1.00 |
| 23 | 8 | 0.89 | 19 | 10 | 1.00 |
| 24 | 8 | 0.89 | 20 | 10 | 1.00 |
| 25 | 8 | 0.89 | 21 | 10 | 1.00 |
| | S-CVI/Ave | 0.90 | | S-FVI/Ave | 0.93 |

**Note**: * I-CVI score ≤ cut-off value, ¥ I-FVI score ≤ cut-off value.

## Psychometric Evaluation

Pilot studies revealed that items about alerts (items 14 and 15) were not answered reliably by almost 25% of participants, since most alerts were switched off as part of the study protocol. These two items were therefore excluded from the EFA, which was performed with a total of 126 questionnaires that had been completely answered. Moreover, through successive iterations of reliability analysis based on corrected total item correlation preceding the EFA, 5 items (items 1, 2, 9, 11 and 16) were sequentially eliminated from the item pool.

Adequacy for EFA of the remaining 14 items (n=126 questionnaires) was confirmed by a KMO value of 0.807 and a significant Bartlett's test (X2= 764; df= 91; p< 0.001). While the Kaiser criterion suggested a four-factor solution (eigenvalues >1), scree plot analysis indicated a clear break point after two factors (Figure 4). Sequential EFA iterations, beginning with four factors, revealed insufficient loadings (|<0.4|) for the third and fourth factors. A three-factor solution was similarly unsuitable due to single-item loading on the third factor. Consequently, a two-factor solution was adopted as the most appropriate factorial structure. Further iterative analysis of factor loading scores resulted in the exclusion of three items (items 4, 8 and 13), that failed to meet the threshold of |>0.4|. Thus, a final EFA with 11 items was performed and revealed a two-factor solution, accounting for 50.3% of the total variance. The first factor, comprising 6 items, was designated as "User acceptance" based on the content of the individual items. The second factor, consisting of 5 factors, was termed "Daily life impact" (Table 2). Cronbach's α of each factor was 0.89 and 0.70, respectively. Subsequently, Cronbach's α of the total scale was 0.84, indicating high internal consistency and good reliability (Table 2) and justifying calculation of a sum score.

## Scoring

Calculating the total score of the 11-item QUSS-CGM resulted in scores ranging from 0 to maximum 44. To facilitate interpretation, scores were transformed to a 0–100 scale. A percentile analysis of the pilot study data based on the distribution of the total scores was conducted to define thresholds for score categorization: scores ≤66 (24th percentile) were classified as low satisfaction, scores between 67 and 91 (25th to 75th percentiles) as average user satisfaction, and scores ≥92 (76th to 100th percentiles) as high user satisfaction (Figure 5). Descriptive statistics reflected the presence of a ceiling effect along the questionnaire, whereas there was no floor effect.

The resulting validated QUSS-CGM, available both in German and English, was established as a two-part questionnaire where the first item should be asked at the beginning of the CGM performance study and the others at the end (Figure 6).

# Discussion

This work aimed at validating the QUSS-CGM, a questionnaire for the assessment of user satisfaction to be used as a standardized tool in CGM performance studies.

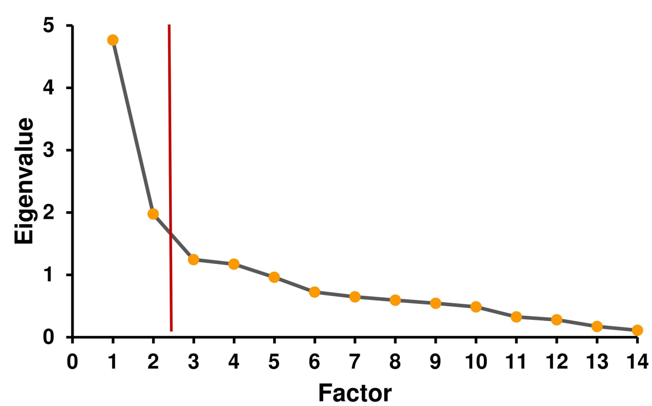


**Figure 4** Scree plot, red line pointing the break (elbow criterion). The scree plot displays the eigenvalues for each extracted factor. A clear inflection point or "elbow" is observed after the second factor, as indicated by the red vertical line. This elbow criterion suggests that a 2-factor solution best represents the underlying dimensional structure of the questionnaire.

**Table 2** Factor Loading Scores from the Two-Factor Solution, Including the Final 11 Items (Designated Accordingly with Keywords) Classified According to Factor 1 "User Acceptance" and Factor 2 "Daily Life Impact"

| | Items | Factor | |
|---|---|---|---|
| | | 1: User Acceptance | 2: Daily Life Impact |
| 3. | Sensor expectations | 0.047 | **0.491** |
| 5. | Wearing comfort | 0.133 | **0.741** |
| 6. | User experience | −0.004 | **0.721** |
| 7. | Everyday convenience | −0.099 | **0.483** |
| 10. | Sensor size | 0.044 | **0.481** |
| 12. | Reliability of readings | **0.676** | −0.063 |
| 17. | Error-free operation | **0.584** | 0.027 |
| 18. | User friendliness | **0.668** | 0.055 |
| 19. | Overall impression | **0.910** | 0.049 |
| 20. | Willingness to reuse | **0.812** | −0.015 |
| 21. | Recommendation to others | **0.941** | −0.001 |
| Cronbach's α | | 0.89 | 0.70 |
| Total Cronbach's α | | 0.84 | |
| Total variance (%) | | 50.3 | |

**Notes**: Principal axis factoring (extraction method) and Oblimin (rotation). Items 4, 8 and 13 removed. Bold text indicates decisive factor loading scores for items' classification.

Pertinent user satisfaction aspects were identified through an extensive literature search of CGM performance studies from the last 20 years. These represented the most frequently addressed factors in such studies, aligning with the intended objective of the QUSS-CGM, to assess aspects that a user can evaluate after an initial, single use of the device.

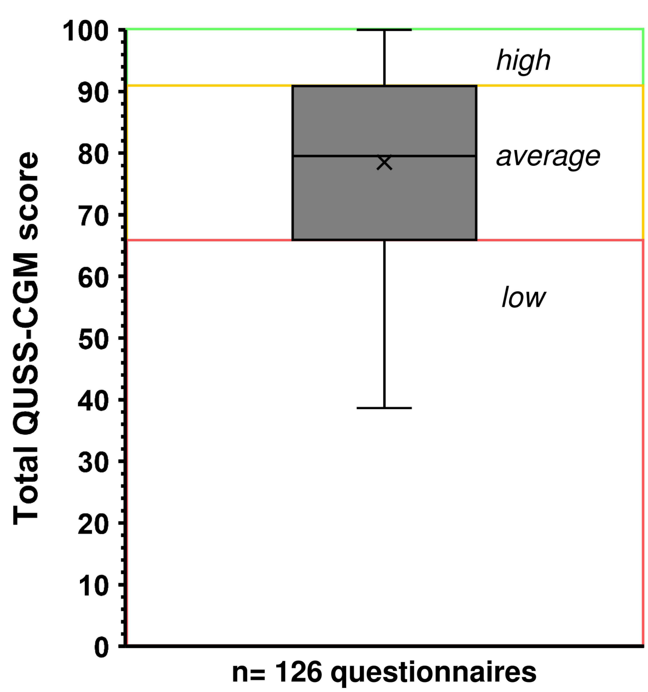


**Figure 5** Total QUSS-CGM score distribution from the final 11-item questionnaire (n=126) and calculated percentiles. Questionnaire for user satisfaction standardized for CGM (QUSS-CGM).

| German | English |
|---|---|
| **Anfang der Studie** | **Beginning of the study** |
| Ich glaube, der Sensor wird mich im Alltag stören. | I think, the sensor will bother me in everyday life. |
| **Ende der Studie** | **End of the study** |
| Der Sensor war angenehm zu tragen. | The sensor was comfortable to wear. |
| Der Sensor hat mich in meinem Alltag beeinträchtigt. | The sensor hindered my daily life. |
| Das Tragen des Sensors war schmerzhaft. | Wearing the sensor was painful. |
| Die Größe des Sensors hat mich in meinem Alltag nicht gestört. | The size of the sensor did not disturb me in my daily life. |
| Ich glaube, die Glukosemessungen waren zuverlässig. | I believe, the glucose measurements were reliable. |
| Das System hat die ganze Zeit einwandfrei funktioniert. | The system was working properly the entire time. |
| Insgesamt empfand ich das CGM-System (Sensor und App) als benutzerfreundlich. | Overall, I perceived the CGM system (sensor and app) as user friendly. |
| Mein Eindruck vom gesamten System (Sensor und App) war positiv. | My impression of the whole system (sensor and app) was positive. |
| Ich würde das System auch in Zukunft nutzen. | I would use the system in future. |
| Ich würde das System anderen Menschen mit Diabetes empfehlen. | I would recommend the system to other people with diabetes. |

| **QUSS-CGM Score** | 0-66 | 67-91 | 92-100 |
|---|---|---|---|
| **User satisfaction** | low | average | high |

**Figure 6** Resulting validated 11-item QUSS-CGM questionnaire (German and English version) and corresponding assessment of user satisfaction based on scoring. Questionnaire for user satisfaction standardized for CGM (QUSS-CGM).

An initial collection of 205 items regarding user satisfaction in CGM performance studies was compiled after an initial systematic literature search, which was organized into a 25-item "draft-questionnaire". The development of the "draft-questionnaire" further continued through content and face validation as described in the material and methods section. Four items addressing pain expectations and ambiguous outcomes (eg, post-sensor removal bleeding) were excluded after experts' content validation based on their potential to introduce response variability and compromise measurement consistency. Items concerning any therapy decisions of the tested CGM systems were also removed since study devices are generally not used for therapy decisions,[36] making it difficult to judge, eg, the accuracy and precision of CGM data. In addition, the experts' opinions on the items about alerts were controversial indicated by an I-CVI below or near the cut-off value, which resulted in one of the three items on this topic being excluded. Assessment of alert-related items is problematic due to default-disabled alerts in CGM performance studies and potential device attribution errors in multi-sensor protocols. Two items on the alerts were classified as initially relevant and left in the questionnaire. Nevertheless, in the pilot studies most alerts were deactivated, thus study participants were not able to rate these items reliably. To guarantee statistical robustness and complete response of the QUSS-CGM, it was decided to exclude these items from the questionnaire. Furthermore, rephrasing some items due to the response process validation was a crucial step since some item elements were recognized as potentially unclear, and by the end of the validation step, efficacy and efficiency were ensured. Overall, the EFA enabled the reduction of the questionnaire to its core 11 components. This reduction eliminated redundancies and enhanced the instrument's efficiency without compromising reliability, resulting in a concise instrument that allows participants to provide reliable responses within a relatively brief timeframe.

The observed ceiling effect in the questionnaire responses during the pilot studies may be attributed to the exclusive evaluation of CE-certified CGM systems, which are expected to yield a high user satisfaction due to the mandatory usability validation required for CE certification. In addition, the questionnaire was designed for between-group comparison rather than discriminative sensitivity analyses. Regardless, the questionnaire proved to be a reliable tool for a thorough and direct assessment of user satisfaction and effectively captured its variations across different CGM systems. In addition, the development of a standardized scoring system facilitates comparative analysis of CGM systems evaluated across independent studies. While the questionnaire did not explicitly address aspects, such as calibration of the system or sensor insertion, these technical and operational factors are likely reflected in the general user experience assessments. Incorporating general user experience measures can provide indirect insight into how technical and operational parameters may impact the overall acceptability from the end-user's perspective.

A strength of this work was the screening based on the systematic literature research on CGM performance evaluations,[2] which allowed for an extensive compilation of study-specific user satisfaction content. Therefore, it served as a foundation comprising all integral aspects to define user satisfaction in the context of CGM performance studies.

Another significant strength was the number of experts that participated, nine, which is nearly the highest recommended number of experts,[12] allowing for a comprehensive quantitative analysis and the number of questionnaires for the EFA (n=126), which exceeds the suggested minimum ratio of respondent to items of 5:1.[17] Moreover, sequential EFA was conducted to optimize item fit and verify sufficient factor loadings, thereby ensuring analytical coherence. A limitation of the study was the number of participants in the response process validation, which was the minimum number recommended,[14] still it was within the acceptable range and the results can be considered adequate.

## Conclusions

This work presents a proposal for the systematic and standardized assessment of user satisfaction in CGM performance studies through validation of the 11-item QUSS-CGM questionnaire. The insights gained can benefit various groups. Patients can provide feedback to guide device design and improve integration with their daily lives, increasing acceptance. Manufacturers can leverage user-centered data to refine product functionality and user experience. Regulatory bodies can incorporate standardized user assessments into a holistic device performance review. By prioritizing the user perspective, CGM systems can be optimized to meet better the needs of individuals managing diabetes.

Since the assessment of user satisfaction is crucial for clinical performance evaluations and patients' acceptance, the QUSS-CGM may serve as a valuable tool as part of a future standardized protocol for CGM performance evaluations.

## Abbreviations

CGM, continuous glucose monitoring; EFA, explorative factor analysis; QUSS-CGM, Questionnaire for User Satisfaction Standardized for CGM performance studies; GME-Q, Glucose Monitoring Experiences Questionnaire; GMSS, Glucose Monitoring Satisfaction Survey; CGM-SAT, CGM Satisfaction scale; GMS, Glucose Monitoring Survey; I-CVI, item-level content validity index; S-CVI/Ave, scale-level content validity index based on the average method; I-FVI, item-level face validity index; S-FVI/Ave, scale-level face validity index based on the average method.

## Ethics Approval and Informed Consent

Approval for the two CGM-Performance studies from which data was collected was obtained by the Ethics Committee of the State Medical Association of Baden-Württemberg (Ethik-Kommission der Landesärztekammer Baden-Württemberg, Stuttgart, Germany), these studies were conducted according to the Declaration of Helsinki. Written informed consent was obtained by all participants prior to enrollment.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Acknowledgments

The authors would like to thank the experts and subjects with diabetes who participated in the validation and the translation of the questionnaire.

## Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

## Disclosure

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: A.B. and M.G.M., M.E., D.W. and C.H. are employees of the Institute for Diabetes Technology Ulm (IfDT), which carries out clinical studies, eg, with medical devices for diabetes therapy on its own initiative and on behalf of various companies. G.F. is the general manager and medical director of IfDT. G.F. and/or its

employees have received speakers' honoraria or consulting fees in the last three years from Abbott, Bionime, BOYDsense, Dexcom, Glucoset, Insulet, i-SENS, Lilly, Lilly Deutschland, Menarini, Novo Nordisk, Perfood, PharmaSens, Roche, Sinocare, Terumo, Vertex, and Ypsomed. D.E. reports Advisory Board member fees from Dexcom Germany and Roche Diabetes Care as well as honoraria for lectures from Berlin-Chemie AG, Sanofi-Aventis, Dexcom Germany, Roche Diabetes Care, Boehringer Ingelheim, and Eli Lilly. The authors report no other conflicts of interest in this work.

The authors would like to disclose that a part of this abstract was presented as a poster at the Deutsche Diabetes Gesellschaft (DDG) congress, held from 8th to 11th May 2024, in Berlin, Germany. The poster's abstract was published in "Abstracts - DDG 2024" in Diabetologie und Stoffwechsel: https://www.thieme-connect.de/products/ejournals/abstract/10.1055/s-0044-1785304.

## References

1. Bailey TS. Clinical implications of accuracy measurements of continuous glucose sensors. *Diabetes Technol Ther*. 2017;19(S2):S51–S54. doi:10.1089/dia.2017.0050
2. Freckmann G, Eichenlaub M, Waldenmaier D, et al. Clinical performance evaluation of continuous glucose monitoring systems: a scoping review and recommendations for reporting. *J Diabetes Sci Technol*. 2023;17(6):1506–1526. doi:10.1177/19322968231190941
3. Freckmann G, Nichols JH, Hinzmann R, et al. Standardization process of continuous glucose monitoring: traceability and performance. *Int J Clin Chem*. 2021;515:5–12. doi:10.1016/j.cca.2020.12.025
4. Eichenlaub M, Pleus S, Rothenbühler M, et al. Comparator data characteristics and testing procedures for the clinical performance evaluation of continuous glucose monitoring systems. *Diabetes Technol Ther*. 2024;26(4):263–275. doi:10.1089/dia.2023.0465
5. Smith MB, Albanese-O'Neill A, Macieira TGR, et al. Human factors associated with continuous glucose monitor use in patients with diabetes: a systematic review. *Diabetes Technol Ther*. 2019;21(10):589–601. doi:10.1089/dia.2019.0136
6. Robinson MA. Using multi-item psychometric scales for research and practice in human resource management. *Human Resour Manage*. 2018;57 (3):739–750. doi:10.1002/hrm.21852
7. Speight J, Holmes-Truscott E, Singh H, Little S, Shaw JAM. Development and psychometric validation of the novel glucose monitoring experiences questionnaire among adults with type 1 diabetes. *Diabetes Technol Ther*. 2019;21(12):691–701. doi:10.1089/dia.2019.0189
8. Polonsky WH, Fisher L, Hessler D, Edelman SV. Development of a new measure for assessing glucose monitoring device-related treatment satisfaction and quality of life. *Diabetes Technol Ther*. 2015;17(9):657–663. doi:10.1089/dia.2014.0417
9. JDRF CGM Study Group. Validation of measures of satisfaction with and impact of continuous and conventional glucose monitoring. *Diabetes Technol Ther*. 2010;12(9):679–684. doi:10.1089/dia.2010.0015
10. Jendrike N, Eichenlaub M, Link M, et al. Comparative performance analysis of manual and updated optional calibration algorithms for the caresens air CGM system. *J Diabetes Sci Technol*. 2025:19322968251351318. doi:10.1177/19322968251351318
11. Eichenlaub M, Waldenmaier D, Wehrstedt S, et al. Performance of three continuous glucose monitoring systems in adults with type 1 diabetes. *J Diabetes Sci Technol*. 2025. doi:10.1177/19322968251315459
12. Yusoff MSB. ABC of content validation and content validity index calculation. *Educ Med J*. 2019;11(2):49–54. doi:10.21315/eimj2019.11.2.6
13. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health*. 2006;29(5):489–497. doi:10.1002/nur.20147
14. Yusoff MSB. ABC of response process validation and face validity index calculation. *Educ Med J*. 2019;11(10.21315).
15. Rahman MA, Yusoff MSB, Roslan NS, Mohammad JA, Ahmad A. Development and validation of the medical professionals resilience scale. *BMC Health Serv Res*. 2021;21(1):482. doi:10.1186/s12913-021-06542-w
16. Williams B, Onsman A, Brown T. Exploratory factor analysis: a five-step guide for novices. *Australasian J Paramed*. 2010;8:1–13. doi:10.33151/ajp.8.3.93
17. Yusoff MSB, Arifin WN, Hadie SNH. ABC of questionnaire development and validation for survey research. *Educ Med J*. 2021;13(1):97–108. doi:10.21315/eimj2021.13.1.10
18. Sachedina N, Pickup JC. Performance assessment of the medtronic-minimed continuous glucose monitoring system and its use for measurement of glycaemic control in type 1 diabetic subjects. *Diabet Med*. 2003;20(12):1012–1015. doi:10.1046/j.1464-5491.2003.01037.x
19. Kubiak T, Wörle B, Kuhr B, et al. Microdialysis-based 48-hour continuous glucose monitoring with GlucoDay: clinical performance and patients' acceptance. *Diabetes Technol Ther*. 2006;8(5):570–575. doi:10.1089/dia.2006.8.570
20. Bode B, Silver M, Weiss R, Martin K. Evaluation of a continuous glucose monitoring system for home-use conditions. *Manag Care*. 2008;17(8):40–45.
21. Meschi F, Bonfanti R, Rigamonti A, et al. Patients' evaluation of nocturnal hypoglycaemia with GlucoDay continuous glucose monitoring in paediatric patients. *Acta Diabetol*. 2010;47(4):295–300. doi:10.1007/s00592-010-0181-9
22. Weinstock RS, Bristol S, Armenia A, Gesswein AC, Bequette BW, Willis JP. Pilot study of a prototype minimally invasive intradermal continuous glucose monitor. *J Diabetes Sci Technol*. 2012;6(6):1454–1463. doi:10.1177/193229681200600627
23. Müller AJ, Knuth M, Nikolaus KS, Krivánek R, Küster F, Hasslacher C. First clinical evaluation of a new percutaneous optical fiber glucose sensor for continuous glucose monitoring in diabetes. *J Diabetes Sci Technol*. 2013;7(1):13–23. doi:10.1177/193229681300700103
24. Bailey TS, Ahmann A, Brazg R, et al. Accuracy and acceptability of the 6-day Enlite continuous subcutaneous glucose sensor. *Diabetes Technol Ther*. 2014;16(5):277–283. doi:10.1089/dia.2013.0222
25. Matuleviciene V, Joseph JI, Andelin M, et al. A clinical trial of the accuracy and treatment experience of the Dexcom G4 sensor (Dexcom G4 system) and enlite sensor (guardian REAL-time system) tested simultaneously in ambulatory patients with type 1 diabetes. *Diabetes Technol Ther*. 2014;16(11):759–767. doi:10.1089/dia.2014.0238

26. Bailey T, Bode BW, Christiansen MP, Klaff LJ, Alva S. The performance and usability of a factory-calibrated flash glucose monitoring system. *Diabetes Technol Ther*. 2015;17(11):787–794. doi:10.1089/dia.2014.0378
27. Nørgaard K, Shin J, Welsh JB, Gjessing H. Performance and acceptability of a combined device for insulin infusion and glucose sensing in the home setting. *J Diabetes Sci Technol*. 2015;9(2):215–220. doi:10.1177/1932296814567326
28. Christiansen MP, Garg SK, Brazg R, et al. Accuracy of a fourth-generation subcutaneous continuous glucose sensor. *Diabetes Technol Ther*. 2017;19(8):446–456. doi:10.1089/dia.2017.0087
29. Edge J, Acerini C, Campbell F, et al. An alternative sensor-based method for glucose monitoring in children and young people with diabetes. *Arch Dis Child*. 2017;102(6):543–549. doi:10.1136/archdischild-2016-311530
30. Ji L, Guo X, Guo L, Ren Q, Yu N, Zhang J. A multicenter evaluation of the performance and usability of a novel glucose monitoring system in chinese adults with diabetes. *J Diabetes Sci Technol*. 2017;11(2):290–295. doi:10.1177/1932296816662884
31. Kropff J, Choudhary P, Neupane S, et al. Accuracy and longevity of an implantable continuous glucose sensor in the PRECISE study: a 180-day, prospective, multicenter, pivotal trial. *Diabetes Care*. 2017;40(1):63–68. doi:10.2337/dc16-1525
32. Olafsdottir AF, Attvall S, Sandgren U, et al. A clinical trial of the accuracy and treatment experience of the flash glucose monitor FreeStyle Libre in adults with type 1 diabetes. *Diabetes Technol Ther*. 2017;19(3):164–172. doi:10.1089/dia.2016.0392
33. Rumpler M, Mader JK, Fischer JP, et al. First application of a transcutaneous optical single-port glucose monitoring device in patients with type 1 diabetes mellitus. *Biosens Bioelectron*. 2017;88:240–248. doi:10.1016/j.bios.2016.08.039
34. Hansen EA, Klee P, Dirlewanger M, et al. Accuracy, satisfaction and usability of a flash glucose monitoring system among children and adolescents with type 1 diabetes attending a summer camp. *Pediatr Diabetes*. 2018;19(7):1276–1284. doi:10.1111/pedi.12723
35. Massa GG, Gys I, Op't Eyndt A, et al. Evaluation of the FreeStyle® Libre flash glucose monitoring system in children and adolescents with type 1 diabetes. *Horm Res Paediatr*. 2018;89(3):189–199. doi:10.1159/000487361
36. Shah VN, Laffel LM, Wadwa RP, Garg SK. Performance of a factory-calibrated real-time continuous glucose monitoring system utilizing an automated sensor applicator. *Diabetes Technol Ther*. 2018;20(6):428–433. doi:10.1089/dia.2018.0143
37. Szadkowska A, Gawrecki A, Michalak A, Zozulinska-Ziolkiewicz D, Fendler W, Mlynarski W. Flash glucose measurements in children with type 1 diabetes in real-life settings: to trust or not to trust? *Diabetes Technol Ther*. 2018;20(1):17–24. doi:10.1089/dia.2017.0287
38. Aronson R, Abitbol A, Tweden KS. First assessment of the performance of an implantable continuous glucose monitoring system through 180 days in a primarily adolescent population with type 1 diabetes. *Diabetes Obes Metab*. 2019;21(7):1689–1694. doi:10.1111/dom.13726
39. Christiansen MP, Klaff LJ, Bailey TS, Brazg R, Carlson G, Tweden KS. A prospective multicenter evaluation of the accuracy and safety of an implanted continuous glucose sensor: the precision study. *Diabetes Technol Ther*. 2019;21(5):231–237. doi:10.1089/dia.2019.0020
40. Freckmann G, Link M, Kamecke U, Haug C, Baumgartner B, Weitgasser R. Performance and usability of three systems for continuous glucose monitoring in direct comparison. *J Diabetes Sci Technol*. 2019;13(5):890–898. doi:10.1177/1932296819826965
41. Ji L, Guo L, Zhang J, Li Y, Chen Z. Multicenter evaluation study comparing a new factory-calibrated real-time continuous glucose monitoring system to existing flash glucose monitoring system. *J Diabetes Sci Technol*. 2023;17(1):208–213. doi:10.1177/19322968211037991
42. Boscari F, Ferretto S, Cavallin F, Fadini GP, Avogaro A, Bruttomesso D. Effectiveness of adding alarms to flash glucose monitoring in adults with type 1 diabetes under routine care. *Acta Diabetol*. 2022;59(7):921–928. doi:10.1007/s00592-022-01884-1
43. Garg SK, Kipnes M, Castorino K, et al. Accuracy and safety of dexcom G7 continuous glucose monitoring in adults with diabetes. *Diabetes Technol Ther*. 2022;24(6):373–380. doi:10.1089/dia.2022.0011
44. Hochfellner DA, Simic A, Taucher MT, et al. Accuracy assessment of the glucomen(®) day CGM system in individuals with type 1 diabetes: a pilot study. *Biosensors*. 2022;12(2). doi:10.3390/bios12020106
45. Laffel LM, Bailey TS, Christiansen MP, Reid JL, Beck SE. Accuracy of a seventh-generation continuous glucose monitoring system in children and adolescents with type 1 diabetes. *J Diabetes Sci Technol*. 2023;17(4):962–967. doi:10.1177/19322968221091816
46. Rebec M, Cai K, Dutt-Ballerstadt R, Anderson E. A prospective multicenter clinical performance evaluation of the C-CGM system. *J Diabetes Sci Technol*. 2022;16(2):390–396. doi:10.1177/1932296820964574
47. Yeoh E, Png D, Khoo J, et al. A head-to-head comparison between guardian connect and FreeStyle Libre systems and an evaluation of user acceptability of sensors in patients with type 1 diabetes. *Diabetes Metab Res Rev*. 2022;38(7):e3560. doi:10.1002/dmrr.3560
48. Alva S, Brazg R, Castorino K, Kipnes M, Liljenquist DR, Liu H. Accuracy of the third generation of a 14-day continuous glucose monitoring system. *Diabetes Therapy*. 2023;14(4):767–776. doi:10.1007/s13300-023-01385-6