Received: 5 December 2025 | Revised: 30 December 2025 | Accepted: 2 January 2026

DOI: 10.1111/dom.70460

REVIEW ARTICLE

WILEY

International clinical opinion on transparency, standardisation, and calibration alignment in the performance evaluation of systems for continuous glucose monitoring

John S. Pemberton BSc[1] | Robert C. Andrews MD[2,3] | Katharine Barnard-Kelly PhD[4] | Tadej Battelino MD[5,6] | Thomas Danne MD[7] | Lutz Heinemann PhD[8] | Partha Kar MD[9,10] | Alistair Lumb MD[11,12] | David M. Maahs MD[13,14] | Julia K. Mader MD[15] | Chantal Mathieu MD[16] | Viswanathan Mohan MD[17,18] | Helen R. Murphy MD[19,20,21] | Eleanor M. Scott MD[22] | Jennifer L. Sherr MD[23] | Carmel E. Smart PhD[24] | Martin Tauschmann MD[25,26] | Amanda Williams MSc[27] | Emma G. Wilmot MD[28,29] | Dessi P. Zaharieva PhD[13,14] | Othmar Moser MD[30,31]

Correspondence
Othmar Moser, Exercise Physiology, Training and Training Therapy Research Group, Institute of Human Movement Science, Sport and Health, University of Graz, Graz, Austria.
Email: Othmar.moser@uni-graz.at

Abstract

Continuous glucose monitoring (CGM) is now central to diabetes management, yet variation in how respective medical products are evaluated limits meaningful comparison between CGM systems. Three barriers currently constrain reliable interpretation of glucose-derived measures. The first is limited transparency: in several regulatory settings, particularly those using Conformité Européenne marking, clinical-study reports, reference-method information and analytical documentation required for market authorisation are not publicly accessible. The second barrier is heterogeneity in study procedures. Existing evaluations use different reference-glucose methods, sampling strategies, glucose-manipulation protocols and participant characteristics, leading to accuracy estimates that cannot be interpreted consistently across systems. The third barrier is calibration alignment. Even with full transparency and aligned procedures, CGM systems may differ because their calibration algorithms are trained on distinct reference-glucose datasets, influencing reported glucose ranges, automated insulin-delivery behaviour and interpretation during device transitions. A modified Delphi process involving clinicians, laboratory scientists, and researchers identified

Abbreviation: A–P–B, Model for CGM calibration alignment: (A) above capillary blood glucose, (P) physiological blood glucose (between capillary and venous), and (B) below venous blood glucose.; AID, Automated insulin delivery; BG, Blood glucose; CE, Conformité Européenne (European conformity marking); CGM, Continuous glucose monitoring; DGR/DGRs, Dynamic glucose region(s); EUDAMED, European Database on Medical Devices; FDA, U.S. Food and Drug Administration; HbA1c, Glycated haemoglobin; IFCC, International Federation of Clinical Chemistry and Laboratory Medicine; iCGM, Integrated continuous glucose monitoring; ISO, International Organisation for Standardisation; MARD, Mean absolute relative difference; MaRoC, Mean absolute rate of change (ΔBG/Δtime); MHRA, Medicines and Healthcare products Regulatory Agency; SMBG, Self-monitoring of blood glucose; TAR, Time above range (>10.0 mmol/L ≈ >180 mg/dL); TBR, Time below range (<3.9 mmol/L ≈ <70 mg/dL); TGA, Therapeutic Goods Administration (Australia); TIR, Time in range (3.9–10.0 mmol/L ≈ 70–180 mg/dL); TITR, Time in tight range (3.9–7.8 mmol/L ≈ 70–140 mg/dL); UK, United Kingdom; US, United States; YSI, Yellow Springs Instrument.

For affiliations refer to page 2561

these issues as the principal determinants of comparability. During this process, the International Federation of Clinical Chemistry and Laboratory Medicine released a validated framework for performance evaluation of CGM systems, providing a unified approach to reference-method selection, dynamic in-clinic testing, and structured reporting. Adoption would reduce procedural variability but does not resolve calibration-alignment differences. This international clinical opinion proposes a pathway towards internationally interpretable CGM evaluation: immediate transparency of clinical evidence, routine declaration of calibration alignment, and progressive adoption of validated standardised procedures. These steps provide a foundation for reliable interpretation and globally comparable assessment of CGM technologies.

1 | INTRODUCTION

Continuous glucose monitoring (CGM) has become foundational to diabetes management. It enables user-directed insulin adjustment, supports automated insulin delivery (AID), and is now recommended as first-line technology for adults, adolescents, and children with type 1 diabetes, with expanding use in insulin-treated people with type 2 diabetes.[1–4] CGM-derived metrics—including time in range (TIR; 3.9–10.0 mmol/L; 70–180 mg/dL) and time in tight range (TITR; 3.9–7.8 mmol/L; 70–140 mg/dL)—are strongly associated with glycated haemoglobin (HbA1c) and long-term complications.[5–8] These relationships underpin contemporary therapeutic targets, including the widely adopted 70% TIR benchmark.[9]

As the number and diversity of CGM systems increase, three barriers now limit meaningful clinical and scientific comparison across different systems.

The first is transparency. Outside the United States (US), clinical-study reports, comparator methodology and analytical dossiers required for market authorisation are rarely openly accessible, making it difficult to verify whether performance data reflect intended user populations.[10–13]

The second issue is heterogeneity in study procedures. Current performance evaluations of CGM systems use different reference methods, sampling densities, glucose-manipulation protocols, CGM–comparator pairing methodology, and participant characteristics.[13] These variations influence accuracy estimates and mean that ostensibly identical CGM metrics may represent different physiological glucose exposures.

The third issue is the absence of a clinical interpretive language that connects technical performance to real-world therapeutic decision-making. Even if study procedures were fully harmonised, clinicians would still lack a standardised way to understand how different CGM systems map glucose levels in interstitial fluid to blood glucose (BG), how these mappings influence time spent within different glucose ranges, and how to interpret systematic differences between CGM systems in day-to-day care. Without such a framework, glucose results from different CGM systems remain difficult to compare for users and clinicians.

This international clinical opinion evaluates these three issues from a clinical perspective. It explains why transparency is essential, how procedural consistency can strengthen comparability, and why a coherent interpretive framework is required to place technical performance in a clinical context. The aim is to provide clinicians, researchers, and health-system leaders with a clear pathway towards interpretable CGM evaluation and communication.

2 | METHODS

This international clinical opinion was developed using a modified Delphi approach.[14] The objective was to determine, from a clinical standpoint, which features of current CGM evaluation procedures most affect comparability between CGM systems. Iterative written exchanges were conducted within a multidisciplinary author group comprising clinicians, laboratory scientists, and researchers with expertise in CGM performance assessment and national implementation of CGM technologies.

During drafting, the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) Working Group on CGM published the first scientifically validated framework for standardising CGM performance evaluation.[15] The group incorporated this framework where it enhanced clinical clarity across the domains identified by the writing group: (i) the transparency of underlying performance data, (ii) the impact of heterogeneous study procedures, and (iii) the need for consistent language describing accuracy, bias and calibration alignment.

3 | TRANSPARENCY

CGM systems are authorised for non-adjunctive use (which enables insulin dosing, driving, hypoglycaemia treatment) from the age of two.[10–13] Given this reach, performance-evaluation data must be accessible in a form that supports meaningful clinical appraisal. Clinicians, reimbursement organisations, and people with diabetes require evidence that allows fair comparison of accuracy, bias, and reliability across CGM systems.

3.1 | Regulatory access to evidence

Table 1 summarises CGM systems authorised in the EU through Conformité Européenne (CE) marking and the US Food and Drug Administration (FDA), including indicated age range, adjunctive status and availability of clinical data. Comparison of these characteristics with publicly accessible evidence demonstrates substantial asymmetry. Several CGM systems hold CE marking for non-adjunctive use in children despite the absence of publicly available paediatric accuracy data. This persists even though the European Database on Medical Devices (EUDAMED), launched in 2017 with an intended public-access deadline of 2021, is not expected to become fully accessible before 2030.[16] In contrast, FDA 510(k) iCGM submissions (three generic study design and 11 performance requirements) and Class III pre-market approvals for CGM systems provide detailed, publicly accessible clinical documentation, including age-specific accuracy data, and comparator methodology.[17]

3.2 | Clinical implications of limited transparency

From a clinical standpoint, market authorisation without corresponding clinical data availability creates a situation in which the perceived risk–benefit balance cannot be independently verified. This has direct consequences. In Italy, a CGM system won a national tender despite the absence of publicly accessible performance data needed to determine whether it had been evaluated across the intended user groups.[11] In the United Kingdom (UK), an AID system is listed on the national pricing framework that underpins the mandated provision of AID therapy for all children with type 1 diabetes, pregnant women and eligible adults.[18] The AID system has incorporated different generations of its own CGM technology authorised as non-adjunctive from 2 years of age, despite no published accuracy data in relevant paediatric populations.[10] No pivotal-trial evidence is available for the safety or effectiveness of the AID system in any age group, which resulted in clinical organisations in the UK advising against its use pending robust, publicly accessible accuracy and safety data.[19,20]

When data is unavailable, clinical risk is unknown, yet authorisation and procurement decisions may imply that evaluation has occurred. In practice, only national regulators, such as the Medicines and Healthcare products Regulatory Agency (MHRA) in the UK, have the authority to obtain these data under CE marking.[16,21] A formal request for the MHRA to obtain accuracy data for the CGM system and safety/effectiveness data for the AID system from 2 years of age was declined, with the agency stating that its data-access mechanisms would be activated if needed for post-market surveillance (correspondence submitted to the editor).

This places the clinical community in a difficult position when reimbursement frameworks adopt technologies without the evidence required to judge safety or appropriateness, and when the underlying data cannot be obtained for independent assessment.

3.3 | Impact on interpretation and comparability

Limited transparency not only constrains safety evaluation but also prevents meaningful comparison between CGM systems. Differences in study procedures and comparator methods influence accuracy metrics,[13,15,22,23] yet without access to protocols and reference-method data it is not possible to determine whether between-device differences reflect methodological artifacts or genuine physiological differences. Full documentation of study procedures is therefore essential for interpreting device performance and applying it appropriately in clinical practice.

3.4 | Practical steps to resolve the transparency deficit

The transparency deficit can be addressed within existing regulatory structures. Publication of core performance information, reference method, sample size, glucose-range distribution, wear-time stability and age-specific accuracy would allow a meaningful interpretation of CGM performance.[13,15,22,23] Therefore, full clinical data should be publicly accessible upon market authorisation.

In Europe, EUDAMED is designed to host such material[16]; in the US, FDA databases already enable near-complete access[17]; and similar mechanisms could be adopted by other international regulators, such as the Therapeutic Goods Administration (TGA) in Australia.[24] Even in the absence of full procedural standardisation, access to this information would allow independent verification of comparator choice, study design, and bias direction across glucose ranges. These steps are essential, feasible, and commensurate with clinical stakes.

4 | LACK OF STUDY-PROCEDURE STANDARDISATION

Differences between iCGM-approved CGM systems will always exist, and their apparent performance can vary substantially across physiological contexts, including those associated with nocturnal compression-related artefact and rapid glucose dynamics during exercise.[25] However, comparability is undermined by substantial heterogeneity in study design, particularly in reference methods, glucose-manipulation protocols, physiological challenge conditions, and reporting.[13] The heterogeneity makes it difficult to distinguish true device-level differences from artefacts of experimental design.[13]

Until recently, attempts to standardise CGM performance evaluation relied mainly on expert consensus rather than validated methodology. International clinical groups recommended adopting and extending the FDA iCGM special controls,[11,12] and a 2020 multinational technical consensus outlined detailed proposals for harmonising CGM performance assessment.[26] However, neither approach underwent prospective validation. In contrast, the recently published IFCC Working Group on CGM guideline provides the first scientifically

TABLE 1 Regulatory status summary of CGM Systems available in the United States and Europe.

Manufacturer	CGM system (clinical data source)	Published age range (yrs)	FDA age (yrs)	FDA non-adjunctive	CE age (yrs)	CE non-adjunctive	Reference method: Capillary (C), Venous (V), Arterialised-Venous (A-V)	Data available to assess bias to reference method by glucose target ranges
FDA authorised—non-adjunctive								
Abbott	FreeStyle Libre 2[53,54]	≥4	≥4	Yes	≥4	Yes	V	Yes
Abbott	FreeStyle Libre 2 Plus[35,55]	≥2	≥2	Yes	≥2	Yes	V	Yes
Abbott	FreeStyle Libre 3[53,54]	≥2	≥2	Yes	≥2	Yes	V	Yes
Abbott	FreeStyle Libre 3 Plus[35,55]	≥2	≥2	Yes	≥2	Yes	V	Yes
Dexcom	Dexcom G6[56–58]	≥2	≥2	Yes	≥2	Yes	A-V	Yes
Dexcom	Dexcom G7[36,59,60]	≥2	≥2	Yes	≥2	Yes	A-V	Yes
Medtronic	Simplera / Simplera Sync[37,61]	≥2	≥2	Yes	≥2	Yes	V	Yes
Senseonics/Ascensia	Eversense E3[62–64]	≥18	≥18	Yes	≥18	Yes	V	Yes
Senseonics/Ascensia	Eversense 365[65,66]	≥18	≥18	Yes	–	–	V	Yes
FDA authorised—adjunctive								
Medtronic	Guardian Sensor 3[67,68]	≥2	≥7	No	≥2	No	V	Yes
Medtronic	Guardian Sensor 4[a]	≥2	≥7	No	≥2	Yes	V	Yes
Senseonics/Ascensia	Eversense[62,63]	≥18	≥18	No	≥18	Yes	V	Yes
CE marking authorised—non-adjunctive (no FDA authorisation)								
AgaMatrix	ALLYcgm[69]	≥18	–	No	≥18	Yes	C	Yes
Dexcom	Dexcom One[56–58]	≥2	–	No	≥2	Yes	A-V	Yes
Dexcom	Dexcom One+[36,59,60]	≥2	–	No	≥2	Yes	A-V	Yes
Medtrum	A6 TouchCare	≥18	–	No	≥2	Yes	V	No
Medtrum	A7+ TouchCare[a]	≥18	–	No	≥2	Yes	V	No
Medtrum	TouchCare Nano A8[a]	≥18	–	No	≥2	Yes	V	No
MicroTech	AiDEX X[a]	≥18	–	No	≥18	Yes	V	No
MicroTech	Linx[a]	≥18	–	No	≥18	Yes	V	No
Roche	Accu-Chek SmartGuide[70]	≥18	–	No	≥18	Yes	C	Yes
Sinocare	iCan i3/i6[71,74]	≥18	–	No	≥2	Yes	V	Yes
Spirit Healthcare	CareSens Air[69]	≥18	–	No	≥18	Yes	C	Yes
WaveForm/Trinity Bio	Cascade CGM[a]	≥18	–	No	≥18	Yes	V	No
CE marking authorised—adjunctive (no FDA authorisation)								
GlucoRx	GlucoRx AiDEX[72]	≥18	–	No	≥14	No	V	No
Infinovo	Glunovo[73]	≥18	–	No	≥18	No	V	No
Nemaura Medical	SugarBEAT[a]	≥18	–	No	≥18	No	V	No

TABLE 1 (Continued)

Manufacturer	CGM system (clinical data source)	Published age range (yrs)	FDA age (yrs)	FDA non-adjunctive	CE age (yrs)	CE non-adjunctive	Reference method: Capillary (C), Venous (V), Arterialised-Venous (A-V)	Data available to assess bias to reference method by glucose target ranges
POCTech	CT-100/CT-100B[a]	≥18	–	No	≥18	No	V	No
Syai Health	Syai Tag[a]	≥18	–	No	≥18	No	V	No
Urathon	Yuwell CT3[a]	≥18	–	No	≥14	No	V	No

[a]Data on file or in the information for use guide.

validated framework for CGM performance evaluation.[15] A full discussion of the guideline is beyond the scope of this statement, but several elements have direct clinical relevance.

4.1 | Comparator choice

The IFCC recommends measuring BG in capillary samples collected every 15 min, using a measurement method with imprecision ≤2.4% and residual bias ≤2.1% after correction to higher-order metrological standards.[15] Arterialised sampling is excluded as non-physiological.[15] A unified reference standard is crucial to ensure that accuracy metrics reflect sensor performance rather than reference-method variability. Capillary glucose has practical and physiological advantages, but the potential use of venous glucose may need to be considered if regulatory jurisdictions do not converge on a single reference. A consensus reference accepted across regulators is preferable, and the optimal choice will become clearer as standardisation efforts advance.

4.2 | Study-participant requirements

The IFCC recommends enrolling at least 100 adults, with ≥80% living with type 1 diabetes, to ensure both adequate sample size and testing under physiologically demanding conditions.[15] People with type 1 diabetes exhibit rapid post-prandial rises, insulin-mediated declines, and wide rate-of-change variability, providing a natural stress test for CGM systems. This requirement does not imply that CGM systems are intended solely for people living with type 1 diabetes; rather, demonstrating performance under the most challenging physiological conditions increases confidence that accuracy will be adequate across broader user groups.

4.3 | Dynamic glucose regions and unified in-clinic testing

The IFCC recommends that all dynamic glucose challenges occur within a session, beginning with a meal to induce a steep rise, followed by an insulin correction at peak glucose to generate a rapid fall.[15] This produces a mean absolute rate of change (MaRoC) >0.06–0.07 mmol/L/min (1.0–1.2 mg/dL/min).[27] This contrasts with most pivotal studies that used slower protocols (e.g., MaRoC ~0.04 mmol/L/min [0.8 mg/dL/min]), which may underestimate dynamic error.[13]

Dynamic Glucose Regions (DGRs) require ≥7.5% of paired data in each key region (BG low, Alert low, Alert high, BG high), ensuring that physiologically important transitions are adequately represented.[15] The clinical question is straightforward: can a CGM system maintain performance during the rapid transitions that most influence therapeutic decisions? Slow-rise or slow-fall protocols [MaRoC <0.04 mmol/L/min (0.8 mg/dL/min)] do not adequately simulate these conditions.[13,27]

4.4 | Structured reporting

The IFCC recommends reporting accuracy by reference-defined DGRs, not by CGM-derived values, to ensure meaningful comparability.[15] Accuracy is recommended to be summarised by rate-of-change bin, sensor day and wear site.[15] Minimum acceptance criteria include, but are not limited to agreement-rate thresholds (AR20 and AR40) calculated within the DGR zones, reflecting accuracy at moments when critical therapeutic decisions occur.[15] These minimum acceptance thresholds correspond to the performance achieved by current market-leading CGM systems that have been used safely for some years in stand-alone and interoperable AID configurations.[15] Furthermore, standardised dynamic testing protocols also provide a foundation for the structured assessment of CGM trend arrows, ensuring that displayed direction and rate of change accurately reflect physiological glucose dynamics. A proposal for the evaluation of trend arrow accuracy exists within the IFCC recommendations.[15] Although standardised evaluation frameworks are initially recommended in adults,[15] assessment in specific populations such as children and pregnancy remains necessary. As underlying glucose physiology is broadly conserved, the same analytical performance principles are expected to apply, although population-specific validation requirements remain to be determined.

Although Mean Absolute Relative Difference (MARD) is commonly reported and understood, global MARD metrics or even glucose range specific MARD's obscure region-specific and dynamic errors that are most relevant to clinical risk[28] MARD values lack clearly defined clinical acceptability thresholds that are made explicit by agreement rate measures. As a result, MARD provides insufficient information to distinguish clinically acceptable from potentially unsafe error profiles for insulin dosing decisions.[28]

From a clinical standpoint, the change of reporting is important to assess performance for real-life CGM use. However, because IFCC methodology is more rigorous than the iCGM's three generic study design requirements, some of the 11 performance metrics will not be met[29]; this does not indicate declining device performance but reflects standardised methodological precision. Once a CGM system meets the IFCC's analytical standards for non-adjunctive use, clinicians need not scrutinise the technical detail. Clinicians, people living with diabetes, and reimbursement organisations can trust that the CGM system meets a rigorously defined, internationally recognised accuracy standard. A comparable approach already operates effectively for self-monitoring blood-glucose systems (SMBG), where International Organisation for Standardisation (ISO) standard compliance and transparent accuracy data form the basis for national reimbursement and tender qualification.[30] Therefore, international adoption of standardised performance-evaluation protocols and agreed minimum criteria for insulin dosing,[15] implemented through a well-recognised framework such as an ISO standard, supports international convergence towards a single, transparent standard for CGM-based insulin dosing, rather than the current landscape of fragmented, heterogeneous, and often opaque regulatory requirements.[10,11]

At present, some reimbursement decisions still appear to treat CE marking as sufficient clinical justification, enabling a heavy focus on cost because performance is implicitly assumed to be comparable across CGM systems.[11,18–20] In response, clinical groups have begun producing comparison charts[31] that show whether performance-evaluation data are publicly accessible and how each device performs against the core criteria recommended since 2020.[26] These charts clarify which CGM systems meet selected 2020 study design recommendations, which have been assessed under less robust study conditions, and where no clinical data are publicly available. This level of transparency is basic, but it offers a practical map of where the risks of using a given CGM system are well characterised and where they are not.

4.5 | Why procedural standardisation alone cannot achieve CGM comparability

Even universal adoption of IFCC requirements[15] cannot eliminate all between-system differences. Two CGM systems may satisfy minimum performance criteria yet still diverge systematically because their calibration algorithms are trained differently. These calibration alignment differences influence time in different glucose ranges, insulin dosing, and AID behaviour. Even with full transparency and rigorous procedural standardisation, a further interpretive layer is required to understand and communicate the systematic differences that remain, by design, between CGM systems.

5 | CALIBRATION ALIGNMENT AND CLINICAL INTERPRETATION

Calibration alignment describes where a CGM system's estimated glucose values sit relative to the BG compartment used during development and performance evaluation. Although CGM systems measure glucose in interstitial fluid, their outputs aim to represent BG exposure through calibration algorithms trained on historical reference glucose datasets. The structure of these datasets and the modelling used to translate in-vitro sensor sensitivity into in-vivo glucose behaviour differ between manufacturers of CGM systems.[32] Consequently, CGM systems that satisfy accepted accuracy criteria may still diverge systematically in absolute BG estimates. This effect is evident in clinical studies demonstrating reproducible positive- and negative-bias profiles to venous glucose values measured with a laboratory method across commonly used factory-calibrated CGM systems.[33,34]

5.1 | Calibration alignment across physiological blood-glucose compartments

Capillary and venous BG differ systematically because of physiological gradients in skin tissue perfusion, nutrient delivery and glucose extraction. In people living with type 1 diabetes, capillary–venous glucose

differences average around 10%, ranging from ~5% at rest to ~30% during rapid post-prandial transitions.[27] During rapid declines from hyperglycaemia, the gradient can reverse, with capillary glucose falling below venous glucose.[27] This capillary–venous gap defines a physiological exposure corridor within peripheral tissues, with venous glucose generally forming the lower bound and capillary glucose the upper bound.

Factory-calibrated CGM systems trained on different reference-glucose datasets (capillary, venous or arterialised-venous) will therefore position their glucose estimates at different points within this physiological corridor.[32] An independent study comparing three widely used factory-calibrated CGM systems head-to-head, using venous glucose and reference-defined glucose ranges, demonstrated reproducible alignment patterns: some CGM systems read above venous glucose (positive bias), while others read below it (negative bias).[27] These directional patterns are also evident, though less pronounced, in pivotal venous-glucose datasets submitted for regulatory authorisation for the same systems.[35–37] The attenuation likely reflects the lower MaRoC in pivotal trials [~0.04 mmol/L/min (0.8 mg/dL/min)][35–37] when compared with the more dynamic conditions of the head-to-head study [0.07 mmol/L/min (1.2 mg/dL/min)],[27] where faster glucose transitions reveal calibration-alignment differences more clearly, particularly during rising glucose levels exceeding 10.0 mmol/L (180 mg/dL) (Figure 1).

These findings indicate that calibration alignment reflects three manufacturer-defined factors[32]: (i) the reference dataset used for algorithm training, (ii) the mapping of in-vitro sensor sensitivity to in-vivo glucose behaviour, and (iii) the adaptive filtering applied on-body. These are systematic properties of each CGM system, not random sensor variability, and they create predictable patterns of over- or under-reading relative to venous glucose. To support clinical interpretation, these patterns require a simple heuristic or model that translates alignment behaviour into therapeutic understanding.

5.2 | Calibration alignment model for clinical interpretation

For clinical communication, alignment patterns can be summarised using the A-P-B model (Figure 2 and Table 2):

- Zone A (Above): CGM BG estimates above the physiological corridor.
- Zone P (Physiological corridor): CGM BG estimates within the venous–capillary glucose corridor.
- Zone B (Below): CGM BG estimates consistently below the physiological corridor.

In a head-to-head study, when compared with venous glucose reference measurements, all three evaluated CGM systems fell within Zone P for TBR but diverged in TIR and TAR.[27] Two systems remained within Zone P across most TIR and TAR, whereas one system predominantly occupied Zone B for TIR and TAR.[27] This explains why the same three CGM systems still differ by up to 10% in TIR or TITR for the same individual, despite exposure to the same physiological conditions.[38]

5.3 | Transparency requirements for calibration alignment

Only a minority of CGM systems have publicly accessible venous-glucose datasets with glucose ranges defined by reference values rather than CGM values, and these are the only datasets sufficient to characterise A-P-B alignment (Table 1). Without declaration of calibration alignment alongside accuracy metrics, it remains difficult to interpret cross-system differences in TIR, TITR, or AID performance, even under standardised analytical conditions.

6 | CLINICAL IMPLICATIONS OF COMPARABILITY GAPS

Without transparency, without standardised study procedures and without a clear language to describe calibration alignment, CGM-derived metrics risk being interpreted in ways that do not reflect true physiological glucose exposure. Across clinical domains, these gaps can lead to misclassification of glycaemic control, inappropriate treatment adjustments, misleading optimisation of automated insulin delivery and incorrect assumptions about physiological change when users transition between systems. The following sections outline the clinical settings in which these issues are most relevant and illustrate why a coherent interpretive layer is needed alongside robust analytical standards.

6.1 | Automated insulin delivery

Systematic reviews AID studies show end of trial TIR and mean sensor glucose values can differ by up to approximately 13% and 1.5 mmol/L (27 mg/dL), respectively, despite corresponding HbA1c differences of no more than 1 mmol/mol (0.1%).[39–41] Algorithmic variation may explain part of the observed divergence in TIR; however, AID algorithms adapt to the glucose values they receive and cannot identify or correct for systematic CGM calibration bias. Consequently, algorithmic differences alone cannot explain persistent discrepancies between CGM-derived TIR or mean sensor glucose and HbA1c. Evidence that identical individuals can exhibit differences exceeding 10% in TIR when wearing different market-leading CGM systems under standardised laboratory conditions,[27,42] as well as during real-world use,[38] further supports calibration alignment is likely an important contributor.

Taken together, these data suggest that differences in calibration alignment may translate into clinically meaningful differences in HbA1c. A clinical example illustrates this limitation. An individual switching an interoperable AID system from a CGM calibrated in Zone B to one calibrated in Zone P may achieve a very similar TIR if the algorithm successfully adapts to the more rapidly changing Zone P

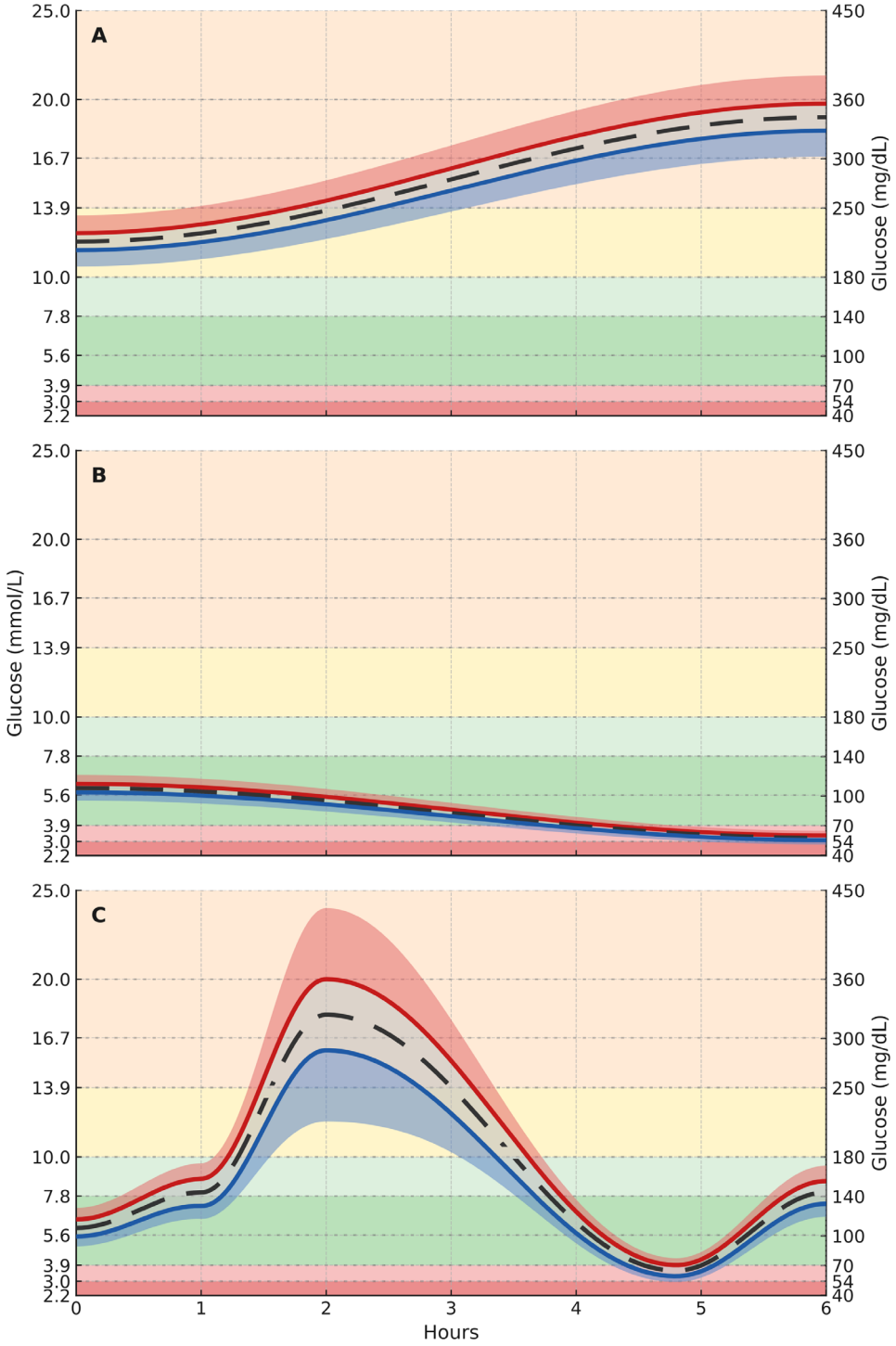

FIGURE 1 In-clinic blood-glucose profiles for different glucose manipulations. The figure contrasts traditional slow-rise (A) and slow-fall (B) protocols with the rapid glycaemic transitions that occur in everyday life (C). A unified in-clinic session combining a postprandial rise with an insulin-driven decline exposes CGM systems to the physiological conditions under which therapeutic decisions are made. This approach enables assessment of CGM performance during the periods when accuracy is most clinically consequential. Colour-coded glucose domains: Orange (TAR Level 2): >13.9 mmol/L (>250 mg/dL); Yellow (TAR Level 1): 10.0–13.9 mmol/L (180–249 mg/dL); Green (TIR): 3.9–10.0 mmol/L (70–180 mg/dL); Dark green (TITR): 3.9–7.8 mmol/L (70–140 mg/dL); Light red (TBR Level 1): 3.0–3.9 mmol/L (54–70 mg/dL); Red (TBR Level 2): <3.0 mmol/L (<54 mg/dL). A-P-B interpretive zones: Zone A (overestimation) lies above capillary blood glucose; Zone P (physiological zone) represents the interval between capillary and venous blood glucose; Zone B (underestimation) lies below venous blood glucose.

glucose values, particularly during hyperglycaemia. If achieved, given the data discussed, this could plausibly result in an apparent HbA1c reduction of approximately 5 mmol/mol (0.5%), despite little change in CGM derived TIR and mean sensor glucose values.

This example highlights the limitations of interpreting CGM-derived metrics in isolation, without matched HbA1c values derived from at least 28 days of CGM data[43] and underscores the importance of calibration alignment when comparing outcomes across systems. For future AID comparisons, standardised CGM performance evaluation will be essential; however, understanding calibration alignment will remain critical for interpreting residual disparities between CGM metrics (TIR and mean sensor glucose) and HbA1c across systems.[38–41]

6.2 | Pregnancy

Pregnancy requires tighter glycaemic targets and features rapid, hormonally driven changes in glucose dynamics. Higher CGM

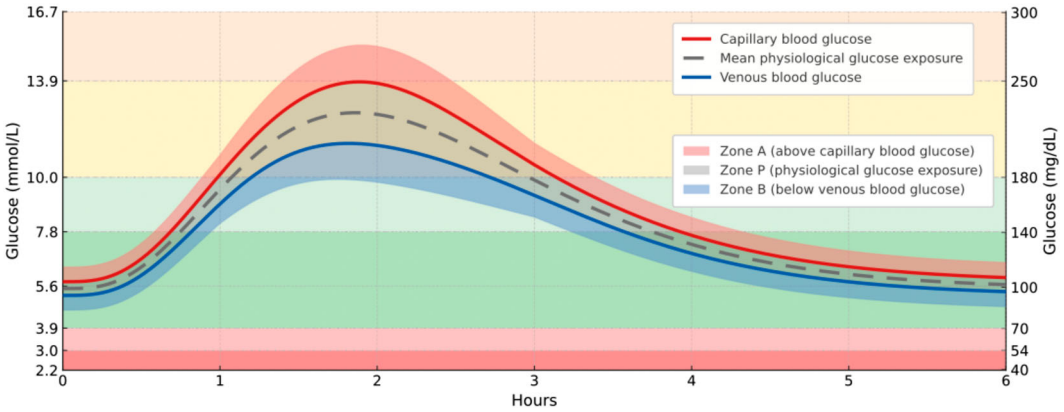

FIGURE 2 Conceptual model of physiological blood-glucose zones. Capillary glucose rises earlier and peaks higher after meals, while venous glucose sits lower and lags during rapid glycaemic transitions. The shaded grey region between the capillary and venous traces represents the physiologically plausible interval within which true blood glucose lies. CGM systems sense interstitial glucose but report an estimate of blood glucose; this interval therefore provides a reference for understanding potential calibration alignment. A-P-B interpretive zones: Zone A (overestimation) lies above capillary blood glucose; Zone P (physiological zone) represents the interval between capillary and venous blood glucose; Zone B (underestimation) lies below venous blood glucose. Colour-coded glucose domains: Orange (TAR Level 2): >13.9 mmol/L (>250 mg/dL); Yellow (TAR Level 1): 10.0–13.9 mmol/L (180–249 mg/dL); Green (TIR): 3.9–10.0 mmol/L (70–180 mg/dL); Dark green (TITR): 3.9–7.8 mmol/L (70–140 mg/dL); Light red (TBR Level 1): 3.0–3.9 mmol/L (54–70 mg/dL); Red (TBR Level 2): <3.0 mmol/L (<54 mg/dL).

TABLE 2 Clinical interpretive model for calibration alignment relative to physiological blood-glucose compartments.

Zone	Clinical interpretation	Average bias of CGM values relative to reference glucose
Zone A	Estimated glucose consistently above physiological blood glucose may prompt earlier or more frequent correction doses and may reduce the apparent exposure to hypoglycaemia.	Above capillary blood glucose
Zone P	Estimated glucose within the physiological corridor reflects alignment with typical peripheral blood-glucose behaviour and corresponds to the range underpinning legacy outcome evidence for CGM devices.	Between capillary and venous blood glucose
Zone B	Estimated glucose consistently below physiological blood glucose may generate more conservative insulin-dosing signals and may under-report hyperglycaemia.	Below venous blood glucose

glucose readings are associated with adverse perinatal outcomes, and AID systems show improved maternal and neonatal outcomes compared to standard care with CGM and insulin pumps and pens.[44,45] Achieving ≥70% time within pregnancy-specific targets (3.5–7.8 mmol/L; 63–140 mg/dL)[46] may be compromised if mean glucose is underestimated because of calibration alignment rather than true physiology.

Evidence for AID use in pregnancy remains limited to a small number of devices, with no controlled, head-to-head comparisons.[44,47] Early real-world data suggest device-specific outcome differences,[48] but these require confirmation. Calibration alignment may therefore carry particular significance in pregnancy, where small changes in mean glucose can influence clinical insulin delivery and clinical decisions.[44–46]

6.3 | Early-stage type 1 diabetes, screening and use beyond diabetes

CGM is increasingly used in early-stage type 1 diabetes,[49,50] pre-diabetes,[51] and in exploratory applications beyond diabetes.[51] In these settings, absolute glucose values hold more interpretive weight than in established insulin therapy. Small deviations from normoglycaemia may prompt diagnostic pathways, risk stratification or behavioural interventions.

Differences in calibration alignment, dynamic behaviour, and reference method can therefore alter whether glycaemic excursions cross clinically important thresholds. Standardised performance evaluation and transparent reporting of alignment help separate genuine physiological patterns from artefact.

6.4 | Research, clinical trials and data science

Usage of CGM is now integral to clinical trials, real-world evidence studies, and machine-learning applications. Many studies pool CGM data across devices or compare interventions using glucose sensors with distinct analytical characteristics.[39–41] Without standardised procedures and clarity about alignment, CGM system-

TABLE 3 Recommended actions to improve the transparency, interpretability and comparability of CGM performance.

Phase	Action	Clinical impact
1. Transparency	Publish full clinical-study reports for all CGM systems, including age ranges, diabetes type, comparator method, sampling density, glucose distributions and analytical procedures, through existing regulatory databases (EUDAMED, FDA, TGA, MHRA).	Enables verification that evidence reflects approved indications (e.g., paediatrics, non-adjunctive use). Supports informed clinical choice, fair procurement and identification of evidence gaps.
	Disclose full performance-study design: reference method, pairing strategy, dynamic-testing procedures.	Allows interpretation of accuracy metrics and separates methodological artefact from true device behaviour.
2. Calibration alignment	Make publicly available all performance-evaluation data comparing CGM estimates with the reference glucose across TBR, TIR and TAR, using reference-defined bins. Report calibration alignment relative to blood-glucose compartments (Zone A, Zone P, Zone B) in regulatory submissions and instructions for use.	Provides necessary context for interpreting glycaemic metrics and supports safe transitions between CGM systems and AID platforms.
3. Procedural standardisation (from IFCC recommendations[15])	Adopt a single traceable comparator meeting international bias and precision criteria, using retrospective calibration when required.	Ensures accuracy metrics reflect sensor performance rather than reference-method variability. Improves comparability across studies and devices.
	Conduct unified in-clinic glucose-challenge sessions achieving physiologically representative rates of change, e.g., MaRoC >0.06–0.07 mmol/L/min (1.0–1.2 mg/dL/min).	Tests CGM function under conditions most relevant to therapeutic risk, including rapid postprandial rises and insulin-driven declines.
	Report performance using IFCC analytical domains (glucose ranges, rate-of-change bins, sensor day, wear site, agreement and precision metrics).	Enables like-for-like comparison and strengthens clinical interpretation of CGM metrics in relation to established outcome evidence.

Abbreviations: AID, automated insulin delivery; CGM, continuous glucose monitoring; EUDAMED, European Database on Medical Devices; FDA, Food and Drug Administration; IFCC, International Federation of Clinical Chemistry and Laboratory Medicine; MaRoC, mean absolute rate of change; TGA, Therapeutic Goods Administration.

specific biases can both obscure and inflate apparent treatment effects.

6.5 | Device transitions and longitudinal care

People may change CGM systems because of procurement decisions, formulary shifts, device upgrades or personal preference. Comparative studies show that different CGM systems can produce different mean glucose and TIR values within the same individual, even when all meet FDA regulatory requirements.[38] Without transparency about alignment and dynamic performance, such changes may be misinterpreted as physiological improvement or deterioration. In clinical practice, individuals using different CGM systems may present with similar CGM-derived metrics yet materially different HbA1c values. While biological variation contributes to this discordance,[52] systematic CGM calibration alignment may also influence estimated mean glucose and warrants consideration.

Standardised evaluation and clear alignment characterisation would allow these transitions to be interpreted with confidence. In the absence of these elements, clinicians and users are left to navigate differences without the information required to judge whether observed changes reflect physiology or device characteristics. This underscores the need for a coherent framework that supports consistent interpretation across CGM systems.

7 | RECOMMENDATIONS

CGM comparability requires three coordinated actions (Table 3). First, full transparency of pivotal-study data; second, declaration of calibration alignment; and third, adoption of IFCC-standardised procedures. Transparency enables independent clinical appraisal and can be implemented immediately through existing regulatory mechanisms. Calibration-alignment declaration provides the interpretive context needed to understand systematic differences between systems, even when accuracy criteria are met. IFCC-standardised procedures establish the validated analytical framework required for future evaluations and harmonised international reporting. Together, these measures provide a practical pathway towards globally interpretable CGM performance data.

8 | CONCLUSION

A coherent pathway is required to achieve comparability of CGM systems and clinically reliable interpretation of CGM data. Immediate transparency allows independent scrutiny; calibration-alignment declaration provides the interpretive context for understanding systematic differences between systems; and adoption of IFCC-standardised procedures will strengthen analytical comparability over time (Figure 3). Together, these steps support meaningful clinical interpretation, improve

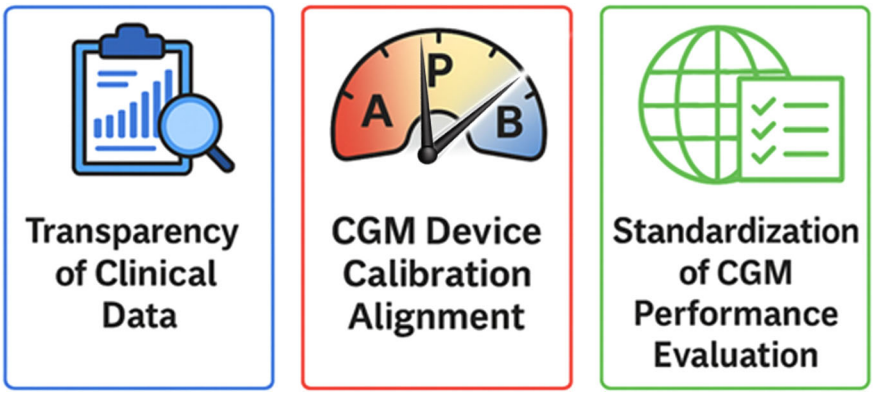

FIGURE 3 The three phases towards global standardisation of accuracy evaluation of CGM systems. Three sequential phases, transparency, calibration alignment, and procedural standardisation, outline a pragmatic pathway towards global comparability of accuracy data obtained with different CGM systems. Calibration alignment is defined using the A–P–B model: Zone A represents glucose estimates above capillary (red = oxygenated) blood glucose; Zone P represents the physiological corridor between capillary and venous blood glucose; and Zone B represents glucose estimates below venous (blue = less oxygenated) blood glucose. The Zone P needle represents a CGM system reading within physiological blood glucose exposure, whereas the Zone B needle represents a CGM system reading below physiological exposure.

therapeutic decision-making for people living with diabetes, and enable internationally comparable evaluation of CGM systems.

AUTHOR CONTRIBUTIONS

All authors contributed to the review and editing of the manuscript and approved the final version. Design: John S. Pemberton, Lutz Heinemann, Julia K. Mader, Tadej Battelino, Chantal Mathieu, Othmar Moser. Conduct/Data Collection: John S. Pemberton, Othmar Moser. Analysis: John S. Pemberton, Othmar Moser. Writing—Manuscript: John S. Pemberton, Robert C. Andrews, Katharine Barnard-Kelly, Tadej Battelino, Thomas Danne, Lutz Heinemann, Partha Kar, Alistair Lumb, David M. Maahs, Julia K. Mader, Chantal Mathieu, Viswanathan Mohan, Helen R. Murphy, Eleanor M. Scott, Jennifer L. Sherr, Carmel E. Smart, Martin Tauschmann, Amanda Williams, Emma G. Wilmot, Dessi P. Zaharevia, Othmar Moser.

AFFILIATIONS

[1]Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK
[2]University of Exeter Medical School, Exeter, UK
[3]NIHR Exeter Clinical Research Facility, Exeter, UK
[4]Barnard Health Research, Southampton, UK
[5]Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia
[6]University Medical Center Ljubljana, Ljubljana, Slovenia
[7]Chief Medical Officer, Breakthrough T1D, Lisbon, New York City, USA
[8]Science Consulting in Diabetes GmbH, Düsseldorf, Germany
[9]Department of Diabetes and Endocrinology, Portsmouth Hospitals NHS University Trust, Portsmouth, UK
[10]NHS England, London, UK
[11]Oxford University Hospitals NHS Foundation Trust, Oxford, UK
[12]Oxford NIHR Biomedical Research Centre, Oxford, UK
[13]Department of Pediatrics, Division of Pediatric Endocrinology, Stanford University, Stanford, California, USA
[14]Stanford Diabetes Research Center, Stanford University, Stanford, California, USA
[15]Division of Endocrinology and Diabetology, Medical University of Graz, Graz, Austria
[16]Endocrinology, UZ Gasthuisberg, KULeuven, Leuven, Belgium
[17]Madras Diabetes Research Foundation (ICMR Collaborating Centre of Excellence), Chennai, India
[18]Dr Mohan's Diabetes Specialities Centre (IDF Centre of Excellence in Diabetes Care), Chennai, India
[19]Norfolk and Norwich University Hospitals NHS Foundation Trust, Norwich, UK
[20]Norwich Clinical Trials Unit, University of East Anglia, Norwich, UK
[21]Norwich Medical School, University of East Anglia, Norwich, UK
[22]Institute of Cardiovascular and Metabolic Medicine, School of Medicine, University of Leeds, Leeds, UK
[23]Yale School of Medicine, New Haven, Connecticut, USA
[24]Department of Endocrinology, John Hunter Children's Hospital, Newcastle, Australia
[25]Institute of Human Movement Science, Sport & Health, Medical University of Graz, Graz, Austria
[26]Division of Paediatric Pulmonology, Allergology and Endocrinology, Medical University of Vienna, Vienna, Austria
[27]East Kent Hospitals University NHS Foundation Trust, Kent, UK
[28]Translational Medical Sciences, School of Medicine, University of Nottingham, Derby, UK
[29]Department of Diabetes, University Hospitals of Derby and Burton NHS Foundation Trust, Derby, UK
[30]Exercise Physiology, Training and Training Therapy Research Group, Institute of Human Movement Science, Sport and Health, University of Graz, Graz, Austria
[31]Cardiometabolic Trials Unit, Medical University of Graz, Graz, Austria

ACKNOWLEDGEMENTS

None. Open Access funding provided by Universitat Graz/KEMÖ.

FUNDING INFORMATION

No specific funding was received for this work.

CONFLICT OF INTEREST STATEMENT

John S. Pemberton—Consultancy fees and speaker honoraria from Abbott, Dexcom, Medtronic, Insulet, and Roche. Member of the IFCC Working Group for CGM. Founder of *The Glucose Never Lies Limited*. Robert C. Andrews—No conflicts of interest related to this work. Katharine Barnard-Kelly—Research funding from Breakthrough T1D, Dexcom, and Abbott; speaker fees from Dexcom and Abbott. CEO and founder of Spotlight-AQ and director of the non-profit WellFight CIC. Tadej Battelino—Advisory panels: Novo Nordisk, Sanofi, Eli Lilly, Boehringer, Medtronic, Abbott, Indigo Diabetes, Roche. Speaker honoraria: Eli Lilly, Novo Nordisk, Medtronic, Abbott, Sanofi, Dexcom, Aventis, AstraZeneca, Roche. Research support: Novo Nordisk, Sanofi, Medtronic, Tandem. Chair International Diabetes Federation Europe. Thomas Danne—has received speaker's honoraria and research support from or has consulted for Abbott, AstraZeneca, Boehringer Ingelheim, Dexcom, Lilly, Medtronic, Novo Nordisk, ProventionBio, Roche, Sanofi, and Vertex Pharmaceuticals and is a shareholder of DreaMed. Chief Medical Officer International, Breakthrough T1D. Lutz Heinemann—Consultant: Abbott, Boydsense, Lifecare (Board), Medtronic EU, Dexcom Germany, Roche, Liom, Perfood. Owner: Profil Institut, Science Consulting, diateam. Partha Kar—Research grant or institutional support: Abbott Diabetes Care, Medtronic, Novo Nordisk, GluSense, Sanofi, Sandoz, Diamyd. Advisory boards: Novo Nordisk, Sanofi, Eli Lilly, Boehringer Ingelheim, Medtronic, DreaMed Diabetes, Bayer Health Care. Speaker honoraria: Eli Lilly, Bayer, Novo Nordisk, Medtronic, Abbott, Sanofi, Roche. Holds stock in DreaMed Diabetes. Alistair Lumb—Personal fees: Insulet, Dexcom, Abbott Diabetes Care, Novo Nordisk, Sanofi, Eli Lilly. Institutional research support: Novo Nordisk, Abbott Diabetes Care. Julia K. Mader—Research collaborations: Roche Diabetes Care, Sanofi-Aventis, Tandem, Viatris. Speaker honoraria: A. Menarini Diagnostics, Abbott Diabetes Care, Dexcom, Eli Lilly, Medtrust, MSD, Novo Nordisk, Roche Diabetes Care, Sanofi, Sinocare, Viatris, Ypsomed. Shareholder: decide Clinical Software GmbH, elyte Diagnostics. Corporate role: Chief Medical Officer (CMO), elyte Diagnostics. David M. Maahs—Research support: NIH, NSF, Breakthrough T1D, Helmsley Charitable Trust. Institutional support: Dexcom, Medtronic, Insulet, Bigfoot, Tandem, Roche. Consulting: Abbott, Sanofi, Eli Lilly, Medtronic, Biospex, Kriya, Enable Biosciences, Aditxt, Helmsley Charitable Trust, LifeScan, MannKind, Novo Nordisk, Insulet, Dompe, Provention Bio. Chantal Mathieu—Advisory panels: Novo Nordisk, Sanofi, Eli Lilly, Novartis, Dexcom, Boehringer Ingelheim, Bayer, Roche, Abbott, Medtronic, Insulet, Biomea Fusion, SAB Bio, Vertex. Speakers bureaus: Novo Nordisk, Sanofi, Eli Lilly, Medtronic, Dexcom, Insulet, Abbott, Vertex, Boehringer Ingelheim. Research support: Dexcom, Novo Nordisk, Sanofi (via KU Leuven). President of EASD and EUDF. Viswanathan Mohan—Consultant, speaker, and recipient of research/educational grants from Abbott, Medtronic, Novo Nordisk, Sanofi, Servier, Boehringer Ingelheim, Eli Lilly, Johnson & Johnson, Lifescan, Roche, MSD, Novartis, Aventis, Bayer, USV, Dr. Reddy's, Sun Pharma, INTAS, Lupin, Glenmark, Zydus, IPCA, Torrent, Cipla, Biocon, Primus, Franco Indian, Wockhardt, Emcure, Mankind, Fourrts, Apex, GSK, and Alembic. Member of the IFCC Working Group for CGM. Helen R. Murphy—Research grant or institutional support: Diabetes UK, NIHR, JDRF, Abbott Diabetes Care, Johnson & Johnson, Medtronic, Dexcom. Advisory panels: Medtronic, Ypsomed, and other diabetes technology companies. Speaker honoraria/travel support: Medtronic, Roche, Sanofi, Novo Nordisk, Ypsomed, Dexcom. Eleanor M. Scott—Honoraria and consultancy: Abbott Diabetes Care, Ypsomed Diabetes Care, Lilly Diabetes Care. Jennifer L. Sherr—Consultant: Abbott Diabetes, Insulet, Medscape, Medtronic Diabetes, Vertex, Ypsomed, Zucara. Advisory boards: Cecelia Health, Insulet, MannKind, Medtronic Diabetes, StartUp Health T1D Moonshot, Vertex. Research support: Abbott Diabetes, Dexcom, JDRF/Breakthrough T1D, Insulet, Medtronic, NIH, Provention Bio. Carmel E. Smart—Speaker Honoria: Medtronic, Roche. Institutional Research support: Medtronic, Sanofi, Novo Nordisk. Martin Tauschmann—Advisory board fees: Abbott Diabetes Care, Sanofi. Speaker honoraria: Eli Lilly, Medtronic, Sanofi, Ypsomed. Amanda Williams—Speaker payments: Roche, Abbott, Dexcom, Mylife, Glooko, Air Liquide, Insulet. Emma G. Wilmot—Honoraria for speaking, advisory board work, or consultancy: Abbott Diabetes Care, AstraZeneca, Dexcom, Insulet, Medtronic, Novo Nordisk, Sanofi, Tandem Diabetes Care. Institutional research support: Medtronic, Novo Nordisk. Dessi P. Zaharevia—has received honoraria for speaking engagements from Ascensia Diabetes Care, Insulet Canada, Dexcom Canada, and Medtronic; is on an advisory board for Dexcom and the Diabetes Research Hub. Othmar Moser—Research collaborations: Abbott, Dexcom, Medtronic, Insulet, Roche.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/dom.70460.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

John S. Pemberton https://orcid.org/0000-0002-0730-7879
Tadej Battelino https://orcid.org/0000-0002-0273-4732
Thomas Danne https://orcid.org/0000-0002-8424-3944
Julia K. Mader https://orcid.org/0000-0001-7854-4233
Chantal Mathieu https://orcid.org/0000-0002-4055-5233
Dessi P. Zaharieva https://orcid.org/0000-0002-9374-8469

REFERENCES

1. American Diabetes Association Professional Practice Committee. 7. Diabetes technology: standards of Care in Diabetes-2024. *Diabetes Care*. 2024;47(Suppl 1):S126-S144. doi:10.2337/DC24-S007
2. de Bock M, Agwu JC, Deabreu M, et al. ISPAD clinical practice consensus guidelines 2024: glycemic targets. *Horm Res Paediatr*. 2024; 97(6):546-554. doi:10.1159/000543266
3. Mohan V, Joshi S, Mithal A, et al. Expert consensus recommendations on time in range for monitoring glucose levels in people with diabetes:

an Indian perspective. *Diabetes Ther*. 2023;14(2):237-249. doi:10.1007/S13300-022-01355-4
4. Tauschman M, Cardona-Hernandez R, Desalvo DJ, et al. International Society for Pediatric and Adolescent Diabetes Clinical Practice Consensus Guidelines 2024 diabetes technologies: glucose monitoring. *Horm Res Paediatr*. 2024;97(6):615-635. doi:10.1159/000543156
5. Beck RW, Bergenstal RM, Cheng P, et al. The relationships between time in range, hyperglycemia metrics, and HbA1c. *J Diabetes Sci Technol*. 2019;13(4):614-626. doi:10.1177/1932296818822496
6. Beck, RW, Raghinaru, D, Calhoun, P, Bergenstal, RM. A comparison of continuous glucose monitoring-measured time-in-range 70–180 mg/-dL versus time-in-tight-range 70–140 mg/dL. *Diabetes Technol Ther* 2024; 26(3):151-1155. doi: 10.1089/dia.2023.0380
7. Shah VN, Kanapka LG, Akturk HK, et al. Time in range is associated with incident diabetic retinopathy in adults with type 1 diabetes: a longitudinal study. *Diabetes Technol Ther*. 2024;26(4):246-251. doi:10.1089/DIA.2023.0486
8. Bergenstal RM, Beck RW, Close KL, et al. Glucose management indicator (GMI): a new term for estimating A1C from continuous glucose monitoring. *Diabetes Care*. 2018;41(11):2275-2280. doi:10.2337/DC18-1581
9. Battelino T, Danne T, Bergenstal RM, et al. Clinical targets for continuous glucose monitoring data interpretation: recommendations from the international consensus on time in range. *Diabetes Care*. 2019;42(8):1593-1603. doi:10.2337/dci19-0028
10. Pemberton JS, Wilmot EG, Barnard-Kelly K, et al. CGM accuracy: contrasting CE marking with the governmental controls of the USA (FDA) and Australia (TGA): a narrative review. *Diabetes Obes Metab*. 2023;25(4):916-939. doi:10.1111/DOM.14962
11. Mathieu C, Irace C, Wilmot EG, et al. Minimum expectations for market authorization of continuous glucose monitoring devices in Europe-'eCGM' compliance status. *Diabetes Obes Metab*. 2025;27(3):1025-1031. doi:10.1111/DOM.16153
12. Scharf M, Feriz K, Yepes C, et al. Consensus statement on standardizing CGM evaluation metrics in Latin America: an expert approach. *Diabetol Metab Syndr*. 2025;17(1):1-16. doi:10.1186/S13098-025-01851-0
13. 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
14. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: how to decide its appropriateness. *World J Methodol*. 2021;11(4):116. doi:10.5662/WJM.V11.I4.116
15. Pleus S, Eichenlaub M, Dabla PK, et al. Clinical assessment and acceptance criteria for continuous glucose monitoring (CGM) system performance: a proposed guideline by the IFCC working group on CGM. *Clin Chim Acta*. 2025;580:120728. doi:10.1016/J.CCA.2025.120728
16. Eur-lex. EUR-Lex-02017R0745-20200424-EN-EUR-Lex. 2021 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02017R0745-20200424
17. Food and Drug Administration. Medical Device Databases | FDA. 2022 https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/medical-device-databases
18. NICE (National Institute for Health and Care Excellence). Technology appraisal guidance [TA943] Hybrid closed loop systems for managing blood glucose levels in type 1 diabetes. 2023 https://www.nice.org.uk/guidance/ta943
19. DTN. DTN statement regarding the use of Medtrum HCL systems | The Association of British Clinical Diabetologists. 2024 https://abcd.care/announcement/dtn-statement-regarding-use-medtrum-hcl-systems
20. BSPED & ACDC. BSPED and ACDC Position statement on Medtrum pump and sensor | BSPED. 2025 https://www.bsped.org.uk/news/item/21251/bsped-and-acdc-position-statement-on-medtrum-pump-and-sensor/
21. Eur-lex. EUR-Lex-01993L0042-20071011-EN-EUR-Lex. 1993 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A01993L0042-20071011
22. 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
23. Freckmann G, Pleus S, Eichenlaub M, et al. Recommendations on the collection of comparator measurement data in the performance evaluation of continuous glucose monitoring systems. *J Diabetes Sci Technol*. 2025;19(4):1072-1081. doi:10.1177/19322968251336221
24. Therapeutic Goods Administration (TGA). TGA Performance Statistics Preport 2020–21. 2020 https://www.tga.gov.au/sites/default/files/annual-performance-statistics-report-july-2020-june-2021.pdf
25. Vaughan N. Meta-analysis of a decade of studies assessing accuracy of Abbott FreeStyle libre continuous glucose monitor. *Diabetes Technology and Obesity Medicine*. 2025;1(1):466-474. doi:10.1177/29986702251390418
26. CLSI. Performance metrics for continuous interstitial glucose monitoring. *CLSI Guideline POCT05*. 2nd ed. Clinical and Laboratory Standards Institute; 2020.
27. 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
28. Heinemann L, Schoemaker M, Schmelzeisen-Redecker G, et al. Benefits and limitations of MARD as a performance parameter for continuous glucose monitoring in the interstitial space. *J Diabetes Sci Technol*. 2020;14(1):135-150. doi:10.1177/1932296819855670
29. Eichenlaub M, Waldenmaier D, Pleus S, Haug C, Brandt D, Freckmann G. Compliance with FDA iCGM special controls is dependent on study design and procedures. *J Diabetes Sci Technol*. 2025;19(3):19322968251329880. doi:10.1177/19322968251329879
30. NHSE. NHS England Commissioning recommendations following the national assessment of blood glucose and ketone meters, testing strips and lancets. 2025 https://www.england.nhs.uk/publication/commissioning-recommendations-blood-glucose-and-ketone-meters-testing-strips-and-lancets/
31. Williams A, Kelly B, Fletcher-Salt T, Pemberton J. Making sense of sensors: evaluating CGM devices for safe and personalised insulin management. *J Diabetes Nurs*. 2025;29(3):JDN378.
32. Hoss U, Budiman ES. Factory-calibrated continuous glucose sensors: the science behind the technology. *Diabetes Technol Ther*. 2017;19(S2):S44-S50. doi:10.1089/DIA.2017.0025
33. Hanson K, Kipnes M, Tran H. Comparison of point accuracy between two widely used continuous glucose monitoring systems. *J Diabetes Sci Technol*. 2024;18(3):598-607. doi:10.1177/19322968231225676
34. Denham D. A head-to-head comparison study of the first-day performance of two factory-calibrated CGM systems. *J Diabetes Sci Technol*. 2020;14(2):493-495. doi:10.1177/1932296819895505
35. CDRH. FreeStyle Libre 2 Flash Glucose Monitoring System, FreeStyle Libre 3 Continuous Glucose Monitoring System, K222447: 510 (k) SUBSTANTIAL EQUIVALENCE. 2024 https://www.accessdata.fda.gov/cdrh_docs/reviews/K222447.pdf
36. CDRH. Dexcom G7 510(k) substantial equivalence determination decision summary. 2023 https://www.accessdata.fda.gov/cdrh_docs/reviews/K213919.pdf
37. CDRH. Premarket Approval (PMA) P160007. 2025 https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=p160007s047
38. Freckmann G, Wehrstedt S, Eichenlaub M, et al. A comparative analysis of glycemic metrics derived from three continuous glucose monitoring systems. *Diabetes Care*. 2025;48(7):1213-1217. doi:10.2337/DC25-0129
39. Di Molfetta S, Di Gioia L, Caruso I, et al. Glycaemic control and variability with different commercially available hybrid closed loop systems in people with type 1 diabetes: a systematic review and

meta-analysis of randomized controlled trials. *Diabetes Obes Metab*. 2026;28(S1):3-12. doi:10.1111/DOM.70150
40. Pöhlmann J, Smith-Palmer J, Serné EH, et al. A systematic literature review and meta-analysis of real-world evidence on commercially available automated insulin delivery systems in people with type 1 diabetes. *Diabetes Obes Metab*. 2026;28(S1):13-34. doi:10.1111/DOM.70161
41. Stahl-Pehe A, Shokri-Mashhadi N, Wirth M, et al. Efficacy of automated insulin delivery systems in people with type 1 diabetes: a systematic review and network meta-analysis of outpatient randomised controlled trials. *EClinicalMedicine*. 2025;82:103190. doi:10.1016/J.ECLINM.2025.103190
42. Sanfilippo S, Thurm U, Renfordt M, et al. Performance of three simultaneously used rtCGM systems around a physically active weekend camp in adults with type 1 diabetes: a prospective lab and real-world study. *Diabetes Obes Metab*. 2026;28(3):1867-1873. doi:10.1111/DOM.70369
43. Tozzo V, Genco M, Omololu SO, et al. Estimating Glycemia from HbA1c and CGM: analysis of accuracy and sources of discrepancy. *Diabetes Care*. 2024;47(3):460-466. doi:10.2337/DC23-1177
44. Lee TTM, Collett C, Bergford S, et al. Automated insulin delivery in women with pregnancy complicated by type 1 diabetes. *N Engl J Med*. 2023;389(17):1566-1578. doi:10.1056/NEJMOA2303911
45. Donovan LE, Lemieux P, Dunlop AD, et al. Closed-loop insulin delivery in type 1 diabetes in pregnancy: the CIRCUIT randomized clinical trial. *JAMA*. 2025;334(24):2176-2185. doi:10.1001/JAMA.2025.19578
46. Committee ADAPP, ElSayed NA, McCoy RG, et al. 15. Management of Diabetes in pregnancy: standards of Care in Diabetes—2025. *Diabetes Care*. 2025;48(Supplement_1):S306–S320. doi:10.2337/DC25-S015
47. Benhalima K, Beunen K, van Wilder N, et al. Comparing advanced hybrid closed loop therapy and standard insulin therapy in pregnant women with type 1 diabetes (CRISTAL): a parallel-group, open-label, randomised controlled trial. *Lancet Diabetes Endocrinol*. 2024;12(6):390-403. doi:10.1016/S2213-8587(24)00089-5
48. Quirós C, Wägner AM, Azriel S, et al. A real-world study comparing advanced hybrid closed-loop systems during pregnancy in women with type 1 diabetes. *Diabetes Technol Ther*. 2025. doi:10.1177/15209156251379488
49. Haller MJ, Bell KJ, Besser REJ, et al. ISPAD clinical practice consensus guidelines 2024: screening, staging, and strategies to preserve Beta-cell function in children and adolescents with type 1 diabetes. *Horm Res Paediatr*. 2024;97(6):529-545. doi:10.1159/000543035
50. Starr L, Dutta S, Danne T, Karpen SR, Hutton C, Kowalski A. The urgent need for breakthrough therapies and a world without type 1 diabetes. *Diabetes Ther*. 2025;16(6):1063. doi:10.1007/S13300-025-01735-6
51. Oganesova Z, Pemberton J, Brown A. Innovative solution or cause for concern? The use of continuous glucose monitors in people not living with diabetes: a narrative review. *Diabet Med*. 2024;41(9):e15369. doi:10.1111/DME.15369
52. Hempe JM, Hsia DS. Variation in the hemoglobin glycation index. *J Diabetes Complications*. 2022;36(7):108223. doi:10.1016/J.JDIACOMP.2022.108223
53. CDRH. FreeStyle Libre 2 510(k) SUBSTANTIAL EQUIVALENCE DETERMINATION DECISION SUMMARY I Background Information: A 510(k) Number FreeStyle Libre 2 Flash Glucose Monitoring System D Regulatory Information Product Code(s) Classification Regulation Section Panel QL. 2022 https://www.accessdata.fda.gov/cdrh_docs/reviews/K193371.pdf
54. 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 Ther*. 2023;14(4):767-776. doi:10.1007/S13300-023-01385-6
55. Alva S, Bhargava A, Bode B, et al. Accuracy of a 15-day factory-calibrated continuous glucose monitoring system with improved sensor design. *J Diabetes Sci Technol*. 2025: 19322968251329364. doi:10.1177/19322968251329364
56. CDRH. Dexcom G6 510(k) Substantial equivalence determination decision memorandum instrument only template. 2022 https://www.accessdata.fda.gov/cdrh_docs/reviews/K183206.pdf
57. Wadwa RP, Laffel LM, Shah VN, Garg SK. Accuracy of a factory-calibrated, real-time continuous glucose monitoring system during 10 days of use in youth and adults with diabetes. *Diabetes Technol Ther*. 2018;20(6):395-402. doi:10.1089/DIA.2018.0150
58. 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
59. 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
60. 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*. 2022;17:962-967. doi:10.1177/19322968221091816
61. Clinicaltrials.gov. Study Details|NCT04436822|Evaluation of Updated Continuous Glucose Monitoring (CGM) Form Factor in Adults, Adolescents and Pediatrics|ClinicalTrials.gov. 2025 https://clinicaltrials.gov/study/NCT04436822
62. Garg SK, Liljenquist D, Bode B, et al. Evaluation of accuracy and safety of the next-generation up to 180-day long-term implantable Eversense continuous glucose monitoring system: the PROMISE study. *Diabetes Technol Ther*. 2022;24(2):84-92. doi:10.1089/dia.2021.0182
63. 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
64. CDRH. Premarket Approval (PMA) P160048. 2025 https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P160048
65. Bailey TS, Liljenquist DR, Denham DS, et al. Evaluation of accuracy and safety of the 365-day implantable Eversense continuous glucose monitoring system: the ENHANCE study. *Diabetes Technol Ther*. 2025;27(5):407-411. doi:10.1089/DIA.2024.0592
66. CDRH. 510(k) Premarket Notification K241335. 2025 https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMN/pmn.cfm?ID=K241335
67. 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
68. Slover RH, Tryggestad JB, Dimeglio LA, et al. Accuracy of a fourth-generation continuous glucose monitoring system in children and adolescents with type 1 diabetes. *Diabetes Technol Ther*. 2018;20(9):576-584. doi:10.1089/dia.2018.0109
69. 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. doi:10.1177/19322968251351318
70. Mader JK, Waldenmaier D, Mueller-Hoffmann W, et al. Performance of a novel continuous glucose monitoring device in people with diabetes. *J Diabetes Sci Technol*. 2024;18(5):1044-1051. doi:10.1177/19322968241267774
71. Yan L, Li Q, Guan Q, et al. Evaluation of the performance and usability of a novel continuous glucose monitoring system. *Int J Diabetes Dev Ctries*. 2023;43(4):551-558. doi:10.1007/S13410-022-01112-0/TABLES/4
72. 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*. 2021;17(1):19322968211037990. doi:10.1177/19322968211037991

73. Meng R, Gu T, Yang F, Liu J, Sun Q, Zhu D. Performance evaluation of the Glunovo® continuous blood glucose monitoring system in Chinese participants with diabetes: a multicenter, self-controlled trial. *Diabetes Ther*. 2021;12(12):3153-3165. doi:10.1007/S13300-021-01171-2
74. Jendrike N, Link M, Öter S, et al. Performance of a New Continuous Glucose Monitoring System in German Adults Living with Diabetes. *Diabetes Ther*. 2025. Dec 29. doi:10.1007/s13300-025-01832-6

How to cite this article: Pemberton JS, Andrews RC, Barnard-Kelly K, et al. International clinical opinion on transparency, standardisation, and calibration alignment in the performance evaluation of systems for continuous glucose monitoring. *Diabetes Obes Metab*. 2026;28(4):2551-2565. doi:10.1111/dom.70460