

Consenso formal de expertos sobre estándares de calidad y recomendaciones en investigación para monitores continuos de glucosa. Metodología de panel Delphi modificado

Formal expert consensus on quality standards and research recommendations for continuous glucose monitors. Modified Delphi panel methodology

EDUARDO MÁRQUEZ-RODRÍGUEZ^{1*}, LUIS A. RAMOS-MÁRQUEZ², CARLOS A. ANTILLÓN-FERREIRA³, LUIS A. PÉREZ- ARREDONDO⁴, MANUEL A. CERVERA-CETINA⁵, JOSÉ A. CETINA-CANTO⁶, RAQUEL N. FARADJI-HAZAN⁷, JUAN C. GARNICA-CUÉLLAR⁸, RAÚL IBARRA-SALCE⁹, FERNANDO J. LAVALLE GONZÁLEZ⁴, HUGO LAVIADA-MOLINA¹⁰, MIGUEL A. MADERO-FERNÁNDEZ DEL CASTILLO¹¹, JOSÉ L. MAYORGA-BUTRÓN¹², LEONARDO G. MANCILLAS-ADAME⁴, ROOPA MEHTA¹⁴, PEDRO MENDOZA-MARTÍNEZ¹⁵, ELIZABETH OLAVIDE-AGUILAR⁴, MARISOL ARROYO-HERNÁNDEZ^{16,17}, LAURA MA. REYES-ORTEGA^{13,16}, MARÍA M. BAQUERO-HOYOS^{13,16}, BERTHA B. MONTAÑO VELÁZQUEZ^{16,18}, LUIS F. ROMERO-MORENO^{16,19}, LILIANA VELAZCO-HIDALGO¹⁹ Y MARIO RODRÍGUEZ-VEGA¹⁶

¹Instituto Jalisciense de Metabolismo, Guadalajara, Jal., México; ²Centro de Obesidad, Diabetes y Tecnologías, Hospital México-Americano, Guadalajara, Jal., México; ³Clínica de Endocrinología, Diabetes y Nutrición, CEDYN, Centro Médico ABC Santa Fe, Ciudad de México, México; ⁴Servicio de Endocrinología, Hospital Universitario Dr. José E. González, Monterrey, N.L., México; ⁵Clínica de Mérida, Mérida, Yuc., México; ⁶Medicina Interna y Endocrinología, Starmédica Mérida, Mérida, Yuc., México; ⁷Endocrinología y Diabetes Clínica EnDi, Centro Médico ABC, Ciudad de México, México; ⁸Endocrinología y Biología de la Reproducción Humana, Centro Médico Nacional 20 de Noviembre, ISSSTE, Ciudad de México, México; ⁹Facultad de Medicina, Unidad Saltillo, Universidad Autónoma de Coahuila, Saltillo, Coah., México; ¹⁰Escuela de Ciencias de la Salud, Universidad Marista de Mérida, Mérida, Yuc., México; ¹¹Facultad de Medicina, Universidad Autónoma de Coahuila, Unidad Laguna, Torreón, Coah., México; ¹²Otorrinolaringología y Cirugía de Cabeza y Cuello, Instituto Nacional de Pediatría, Ciudad de México, México; ¹³Facultad de Medicina, Universidad Nacional Autónoma de México, Ciudad de México, México; ¹⁴Departamento de Endocrinología y Metabolismo, Instituto Nacional de Ciencias Médicas y Nutrición, Ciudad de México, México; ¹⁵Departamento de Endocrinología, Centro Médico La Raza, IMSS, Ciudad de México, México; ¹⁶Ibero American Agency for Development & Assessment of Health Technologies (A2DAHT), Ciudad de México, México; ¹⁷Servicio de Neumología, Instituto Nacional de Cancerología, Ciudad de México, México; ¹⁸Hospital de Especialidades, Centro Médico Nacional La Raza, Ciudad de México, México; ¹⁹Fundación Hospital de la Misericordia, Bogotá, Colombia

*Correspondencia:

Eduardo Márquez-Rodríguez

E-mail: doctormarquezr1@hotmail.com

Fecha de recepción: 12-07-2026

Fecha de aceptación: 05-08-2026

DOI: 10.24875/RME.M26000045

Disponible en internet: 28-09-2026

Rev Mex Endocrinol Metab Nutr. 2026;13(SUPL 1):1-36

2462-4144 / © 2026 Sociedad Mexicana de Nutrición y Endocrinología, AC. Publicado por Permanyer. Este es un artículo *open access* bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

RESUMEN

El monitoreo continuo de glucosa (MCG) se ha consolidado como una tecnología central para optimizar el control glucémico y reducir el riesgo de hipoglucemia en múltiples poblaciones con diabetes. Sin embargo, la utilidad clínica y la seguridad del MCG dependen críticamente de la calidad metrológica del sistema, de la forma en que se obtienen y comparan los datos frente a métodos de referencia, de la estandarización de métricas y reportes, y del cumplimiento de estándares regulatorios que garanticen desempeño consistente en condiciones reales de uso. Esta iniciativa presenta el desarrollo de un Consenso Formal de Expertos impulsado por la Sociedad Mexicana de Nutrición y Endocrinología (SMNE), orientado a establecer recomendaciones prácticas para el uso clínico, la evaluación del desempeño y el marco de seguridad del MCG, con énfasis en la comparabilidad de resultados, la generación e integridad de reportes, la gestión de información y la atención al usuario. Se llevó a cabo una búsqueda estructurada en PubMed/Medline empleando estrategias reproducibles que combinaron términos MeSH y palabras clave, priorizando revisiones sistemáticas, metaanálisis, estudios de evaluación de desempeño clínico, consensos internacionales y documentos regulatorios relevantes. La síntesis de evidencia se tradujo en enunciados operativos sometidos a un Delphi modificado, con votación anónima y retroalimentación iterativa, utilizando una escala de Likert de 1 a 9 para cuantificar acuerdo, incertidumbre o desacuerdo, y para refinar definiciones y recomendaciones hasta alcanzar consenso cuando fue posible. Se incorporaron consideraciones metodológicas sobre el comparador de referencia (matriz, compartimento, trazabilidad, control preanalítico-analítico-postanalítico y alineación temporal), así como criterios para asegurar representatividad clínica mediante distribución adecuada de concentraciones de glucosa y tasas de cambio (RoC), evitando sesgos que subestimen el error en extremos clínicamente relevantes. El consenso resultante propone un marco de recomendaciones que integra: 1) métricas mínimas y complementarias para evaluar exactitud numérica, exactitud clínica, precisión, desempeño dinámico (tendencia/lag) y confiabilidad técnica (*gaps* y cobertura); 2) criterios para selección y uso del MCG en investigación clínica, con énfasis en estandarización de procedimientos, transparencia de reporte y comparabilidad entre estudios, y 3) elementos de seguridad centrados en el usuario, incluyendo verificación con medición capilar cuando esté indicado, manejo de discrepancias clínicas, entrenamiento y soporte, y alineación con estándares internacionales y expectativas regulatorias, particularmente los requisitos asociados a sistemas interoperables (iCGM). Este documento busca servir como referencia metodológica y clínica para México, promoviendo el uso seguro y evaluaciones de desempeño comparables del MCG en entornos ambulatorios y en escenarios clínicos relevantes.

Palabras clave: Monitoreo continuo de glucosa. Exactitud. MARD. iCGM. Consenso Delphi. Estándares de desempeño.

ABSTRACT

Continuous glucose monitoring (CGM) has become a central technology for optimizing glycemic control and reducing the risk of hypoglycemia across multiple populations with diabetes. However, the clinical utility and safety of CGM critically depend on the metrological quality of the system; the way data are obtained and compared against reference methods, the standardization of metrics and reports, and compliance with regulatory standards that ensure consistent performance under real-world conditions of use. This document presents the development of a Formal Expert Consensus promoted by the Mexican Society of Nutrition and Endocrinology (SMNE), aimed at establishing practical recommendations for clinical use, performance evaluation, and the safety framework for CGM, with emphasis on comparability of results, report generation and integrity, information management, and user support. A structured search in PubMed/Medline was conducted using reproducible strategies combining MeSH terms and keywords, prioritizing systematic reviews, meta-analyses, clinical performance evaluation studies, international consensus statements, and relevant regulatory documents. The evidence synthesis was translated into operational statements submitted to a modified Delphi process, with anonymous voting and iterative feedback, using a 1-9 Likert scale to quantify agreement, uncertainty, or disagreement, and to refine definitions and recommendations until consensus was reached when possible. Methodological considerations regarding the reference comparator were incorporated (matrix, compartment, traceability, pre-analytical-analytical-post-analytical control, and temporal alignment), as well as criteria to ensure clinical representativeness through an adequate distribution of glucose concentrations and rates of change (RoC), avoiding biases that underestimate error at clinically relevant extremes. The resulting consensus proposes a framework of recommendations integrating: 1) minimum and complementary metrics to assess numerical accuracy, clinical accuracy, precision, dynamic performance (trend/lag), and technical reliability (gaps and coverage); 2) criteria for CGM selection and use in clinical research, emphasizing standardization of procedures, transparency of reporting, and comparability across studies, and 3) user-centered safety elements, including verification with capillary measurement when indicated, management of clinical discrepancies, training and support, and alignment with international standards and regulatory expectations, particularly requirements associated with interoperable systems (iCGM). This document aims to serve as a methodological and clinical reference for Mexico, promoting safe use and comparable performance evaluations of CGM in ambulatory settings and clinically relevant scenarios.

Keywords: Continuous glucose monitoring. Accuracy. MARD. iCGM. Delphi consensus. Performance standards.

GLOSARIO DE TÉRMINOS

Término	Definición operacional
Adjuntivo (<i>adjunctive CGM</i>)	Sistema de MCG cuyo uso para decisiones terapéuticas (p. ej., ajuste de dosis de insulina o tratamiento de hipoglucemia) requiere confirmación con medición capilar/medidor de glucosa en sangre (BGM).
No adjuntivo (<i>non-adjunctive CGM</i>)	Sistema de MCG con desempeño/precisión considerados suficientes para permitir decisiones terapéuticas sin confirmación sistemática con BGM, conforme a indicación y etiquetado del fabricante y la autoridad regulatoria aplicable.
AGP (<i>ambulatory glucose profile</i> , perfil ambulatorio de glucosa)	Representación estandarizada de datos de glucosa (típicamente 10-14 días) que sintetiza mediana, percentiles y patrones diarios para facilitar interpretación clínica y comparabilidad.
AID (<i>automated insulin delivery</i> , administración automatizada de insulina)	Sistemas que combinan MCG con bombas y algoritmos para ajustar la infusión de insulina de forma automatizada (parcial o total), dependientes de transmisión fiable y continua de datos del sensor.
Análisis preanalítico	Etapas que incluye la obtención de la muestra (p. ej., capilar/venosa), tipo de tubo/ anticoagulante, transporte, temperatura y tiempo hasta el análisis; es determinante para la validez de las mediciones comparadoras (referencia) en estudios de desempeño.
Arterialización (venosa arterializada)	Procedimiento (p. ej., calentamiento local) para aproximar características de sangre arterial en una muestra venosa, con implicaciones sobre el sesgo capilar-venoso y, por tanto, sobre estimaciones de desempeño del MCG frente a referencia.
BGM/SMBG (<i>blood glucose monitoring/ self-monitoring of blood glucose</i> , autocontrol capilar)	Medición puntual de glucosa en sangre (generalmente capilar) mediante glucómetro. Puede ser comparador en estudios o herramienta de verificación clínica en dispositivos adjuntivos.
Brecha de disponibilidad de datos (<i>data gaps</i>)	Intervalos clínicamente relevantes sin lecturas del sensor que pueden comprometer seguridad (p. ej., alarmas) y desempeño de sistemas conectados (AID).
Calibración (manual capilar)	Procedimiento mediante el cual el usuario introduce valores de glucosa en sangre (usualmente capilar) para ajustar el algoritmo del sensor; su frecuencia, oportunidad y calidad pueden modificar sesgo e imprecisión del MCG, especialmente en sensores no calibrados de fábrica.
Calibración de fábrica (<i>factory-calibrated</i>)	Configuración en la cual el sensor no requiere calibraciones manuales habituales del usuario; el desempeño depende del control de manufactura, del algoritmo y de la estabilidad durante la vida útil del sensor.
Coefficiente de variación (CV)	Métrica principal estandarizada para cuantificar la variabilidad glucémica. Se calcula dividiendo la desviación estándar entre la glucosa media, expresándose en porcentaje. Clínicamente, un $CV \leq 36\%$ es el objetivo recomendado para definir una glucemia estable y un menor riesgo de hipoglucemia.
Comparador (referencia)	Medición «de referencia» usada para estimar el desempeño del MCG; puede provenir de métodos de laboratorio aceptados o de medición capilar, según el objetivo del estudio. La elección del compartimento (capilar/venoso/arterializado), matriz (sangre/plasma) y método analítico afecta estimaciones de exactitud/precisión.
Comparabilidad entre estudios (<i>comparability</i>)	Grado en que los resultados de desempeño pueden compararse válidamente entre estudios; depende de armonización del comparador, emparejamiento temporal, distribución de glucosas, RoC y reporte transparente de métodos.
Concordancia/ <i>agreement rate</i> (tasa de concordancia)	Proporción de pares MCG-referencia que caen dentro de umbrales predefinidos (p. ej., ± 15 mg/dl o $\pm 15\%$); en marcos regulatorios puede requerirse demostrar cumplimiento usando límites inferiores de confianza.

(Continúa)

GLOSARIO DE TÉRMINOS (CONTINUACIÓN)

Término	Definición operacional
Confianza (LCL, <i>lower confidence limit</i>)	Límite inferior (usualmente unilateral) de un intervalo de confianza usado para demostrar que una tasa de concordancia o un criterio de desempeño cumple un umbral mínimo con un margen estadístico preespecificado.
<i>Consensus error grid</i> (CEG, rejilla de error por consenso)	Método de exactitud clínica que clasifica la discrepancia MCG-referencia por zonas de riesgo clínico (de mínima a peligrosa) para ponderar la relevancia clínica del error más allá del error numérico.
Datos pareados (<i>paired readings</i>)	Observaciones emparejadas MCG-referencia definidas por reglas explícitas de sincronización temporal y ventana de emparejamiento.
Degradación de desempeño (<i>drift</i>)	Cambio sistemático del sesgo/imprecisión del sensor a lo largo de la vida útil (<i>wear period</i>).
Dispositivo MCG (monitoreo continuo de glucosa)	Tecnología que mide glucosa intersticial de forma continua o frecuente por un periodo definido, proporcionando valores y tendencias para apoyar decisiones clínicas y autocuidado.
Distribución de glucosas (<i>glucose distribution</i>)	Proporción de pares en hipoglucemia, euglucemia e hiperglucemia; su balance es crítico para que métricas globales (p. ej., MARD) no oculten desempeño en extremos clínicamente relevantes.
Exactitud (<i>accuracy</i> , exactitud analítica/numérica)	Cercanía del valor del MCG al valor de referencia; se resume con métricas como MARD o tasas de concordancia dentro de umbrales.
Exactitud clínica (<i>clinical accuracy</i>)	Grado en que el error del MCG podría conducir a decisiones clínicas seguras o inseguras; se evalúa con rejillas de error y otros enfoques clínicos complementarios.
Frecuencia de muestreo de referencia (<i>sampling schedule</i>)	Calendario de obtención de mediciones comparadoras durante sesiones en clínica y/o vida libre; determina la capacidad de capturar variabilidad y cambios rápidos, afectando estimaciones de desempeño y de RoC.
Indicador de manejo de glucosa (GMI)	Valor porcentual que estima la hemoglobina glucosilada (HbA1c) de laboratorio basándose en la glucosa media medida por el MCG (idealmente de 14 días o más). Ayuda a orientar las decisiones terapéuticas, aunque clínicamente puede diferir de la HbA1c real debido a variaciones biológicas individuales en la glucosilación de los glóbulos rojos.
Heterogeneidad metodológica	Variación entre estudios en diseño, comparador, emparejamiento, distribución de glucosa y análisis estadístico; limita comparaciones directas entre sistemas y la interpretación de desempeño.
Hipoglucemia (niveles)	Clasificación clínica que suele distinguir: nivel 1 (< 70 mg/dl y $\geq$ 54 mg/dl), nivel 2 (< 54 mg/dl) y nivel 3 (evento grave con alteración funcional y necesidad de asistencia).
iCGM (<i>integrated continuous glucose monitor</i>)	Categoría regulatoria (FDA) para MCG con requisitos especiales de desempeño e interoperabilidad; incluye criterios explícitos para exactitud por rangos y desempeño de tendencia (RoC), además de expectativas de seguridad y documentación.
Interferentes (<i>interfering substances</i>)	Sustancias endógenas o exógenas (p. ej., fármacos, metabolitos) que pueden alterar la lectura del sensor.
Imprecisión (<i>precision</i>)	Grado de dispersión de mediciones repetidas (o entre sensores) bajo condiciones similares; complementa la exactitud.
Lote de fabricación del sensor (<i>sensor lot</i>)	Conjunto de sensores producidos bajo condiciones similares; evaluar múltiples lotes reduce el riesgo de sobreestimar desempeño por selección de producción atípica.
MBG (<i>mean blood glucose</i> , glucosa promedio)	Promedio de lecturas (generalmente del MCG) durante un periodo definido; métrica descriptiva complementaria.

(Continúa)

GLOSARIO DE TÉRMINOS (CONTINUACIÓN)

Término	Definición operacional
MARD (<i>mean absolute relative difference</i> , diferencia relativa absoluta media)	Promedio de $ MCG - \text{referencia} / \text{referencia} \times 100$ en pares de medición; resume exactitud numérica global, pero es sensible al comparador, distribución de glucosa, RoC, emparejamiento y población.
Matriz (sangre total/plasma/suero)	Medio biológico donde se mide glucosa de referencia; la matriz y su conversión (p. ej., sangre total a equivalente de plasma) influyen en sesgo y comparabilidad.
MCG en tiempo real (rtCGM)	MCG que transmite lecturas automáticamente (sin necesidad de escaneo), generalmente con alertas/alarma.
MCG de escaneo intermitente (isCGM/flash)	MCG que requiere «escaneo» para visualizar lecturas; su arquitectura puede influir en disponibilidad de datos y experiencia del usuario.
Periodo de uso del sensor (<i>sensor wear period</i>)	Duración reclamada/indicada durante la cual el sensor debe mantener desempeño estable; las evaluaciones deben cubrir el periodo completo para detectar <i>drift</i> .
Precisión de tendencia (<i>trend accuracy</i>)	Exactitud en la representación del cambio de glucosa en el tiempo (tendencia), evaluada por comparación de RoC del MCG vs. RoC de referencia.
Población de uso intencionado (<i>intended use population</i>)	Grupo clínico para el cual se declara indicación del dispositivo; las evaluaciones deben demostrar desempeño robusto y representativo en dicha población.
Postanalítico	Eta posterior a la medición de referencia que incluye registro, validación, conversión (si aplica), emparejamiento con MCG y reglas de exclusión; influye en trazabilidad y comparabilidad.
Rango glucémico (<i>glucose range</i>)	Segmentación operativa de glucosas (p. ej., < 70, 70-180, > 180 mg/dl) usada para reportar desempeño por estratos.
Referencia trazable (<i>traceability</i>)	Propiedad por la cual el método comparador puede vincularse metrológicamente a estándares de orden superior, reduciendo sesgos sistemáticos.
RoC/ROC (<i>rate of change</i> , tasa de cambio)	Velocidad de cambio de glucosa (mg/dl/min). Es clave para evaluar seguridad durante ascensos/descensos rápidos.
Seguridad del usuario (<i>user safety</i>)	Conjunto de propiedades del sistema (exactitud, alertas, disponibilidad, mitigación de interferentes y usabilidad) orientadas a minimizar daños por mediciones erróneas o fallas del sistema.
Sesgo (<i>bias</i>)	Desviación sistemática del MCG respecto a referencia. Puede variar según compartimento de muestreo, matriz, RoC, tiempo de uso del sensor y condiciones fisiológicas.
<i>Sensor-to-sensor variability</i> (variabilidad entre sensores)	Diferencias de desempeño entre sensores de un mismo sistema; parte de la caracterización de precisión y confiabilidad técnica.
TAR (<i>time above range</i> , tiempo arriba de rango)	Porcentaje de tiempo con glucosa por encima de un umbral (frecuentemente > 180 mg/dl), con posibles estratos adicionales (p. ej., > 250 mg/dl).
TBR (<i>time below range</i> , tiempo abajo de rango)	Porcentaje de tiempo con glucosa por debajo de umbral (frecuentemente < 70 mg/dl), con estratos (p. ej., < 54 mg/dl).
TIR (<i>time in range</i> , tiempo en rango)	Porcentaje de tiempo con glucosa dentro de un rango objetivo (típicamente 70-180 mg/dl).
Usabilidad (<i>usability</i>)	Capacidad del usuario previsto para operar el dispositivo de manera segura y obtener el desempeño esperado; incluye comprensión de alertas, manejo de fallas y acciones ante discrepancias.

(Continúa)

GLOSARIO DE TÉRMINOS (CONTINUACIÓN)

Término	Definición operacional
Validez (<i>validity</i>)	Grado en que el desempeño observado del MCG refleja su desempeño real sin distorsión por comparador, emparejamiento, muestreo o análisis.
Vida útil reclamada (<i>claimed sensor life</i>)	Duración declarada por el fabricante para uso del sensor, para la cual debe demostrarse desempeño consistente.
Ventana de emparejamiento temporal (<i>pairing window</i>)	Intervalo permitido para vincular una lectura de MCG con una medición de referencia, considerando sincronización de reloj y retardo fisiológico (intersticial vs. sangre).

INTRODUCCIÓN

La diabetes sigue siendo una de las condiciones más apremiantes para la salud pública en el mundo, debido a su elevada prevalencia y a su contribución en la mortalidad prematura, la discapacidad y la elevada carga de enfermedad¹. En el año 2021, se reportó que 529 millones de personas vivían con diabetes y se proyecta que para 2050 la cifra superará los mil trescientos millones de personas. De estos pacientes, el 96% de los casos corresponde a diabetes tipo 2 (DT2), lo que señala una desproporción importante de esta entidad a nivel mundial¹. Las tendencias más recientes confirman un incremento sostenido tanto en la prevalencia como en las brechas de tratamiento y control. La colaboración NCD-RisC, a partir de 1,108 estudios poblacionales (141 millones de participantes), reportó que la prevalencia de diabetes se duplicó en adultos desde 1990 y que millones permanecen sin tratamiento, lo que resalta la presencia de rezagos en atención médica y prevención secundaria².

La DT2 contribuye a una mayor proporción del impacto de la enfermedad asociado a la diabetes. En el análisis GBD (*Global Burden of Disease*) 2021, el 52.2% de los años de vida ajustados por discapacidad (DALY) por DT2 se atribuyó al índice de masa corporal (IMC) elevado, lo que indica el rol dominante de la obesidad y del entorno obesogénico en la génesis y progresión de la enfermedad; dicho estudio también sitúa a la diabetes

entre las principales causas de años vividos con discapacidad y de muerte prematura, con marcadas desigualdades regionales¹.

Por el contrario, la diabetes tipo 1 (DT1) es mucho menos prevalente; sin embargo, es más compleja en su manejo y está asociada a descontrol y complicaciones prematuras. En regiones menos desarrolladas, el acceso deficiente a medicamentos y el monitoreo adecuado contribuyen a la complejidad de esta enfermedad³.

Por su parte, la diabetes gestacional (DG) también contribuye de manera sustancial a la carga metabólica materno-infantil. Una revisión publicada en 2024 sintetizó que alrededor del 14% de los embarazos globales cursa con DG, con variabilidad según criterios diagnósticos y riesgos poblacionales, y con tendencia ascendente, paralela a la epidemia de obesidad y DT2. Una estimación estandarizada para los criterios de la Asociación Internacional de los Grupos de Estudio sobre Diabetes y Embarazo (IADPSG, *International Association of the Diabetes and Pregnancy Study Groups*) reportó prevalencias globales y regionales elevadas⁴.

Las cifras en nuestro país también denotan una gran carga de la enfermedad, además de un pobre control de estas. Reportes recientes destacan una prevalencia combinada de diabetes (diagnosticada y no diagnosticada) que se sitúa entre el 15 y 17% (ENSANUT 2018-2020), con un control glucémico subóptimo: solo el 42.1% de

los adultos con diabetes alcanzó un porcentaje de hemoglobina glucosilada (HbA1c) < 7% en 2018 y 39.0% en 2020⁵.

Por otro lado, las cifras actualizadas de ENSANUT 2022 muestran un aumento de la prevalencia nacional del 14.4% (2006) al 18.3% (2022) en paralelo a un incremento de la incidencia documentado por el GBD (de 367 a 496/100,000 entre 2006-2019). Además, el 22.1% de los adultos presentó prediabetes (intervalo de confianza del 95% [IC 95%]: 19.6-24.7), con un gradiente ascendente con la edad (32.8% en ≥ 60 años)⁶. Este reporte también evidencia que el 31.7% de las personas con diabetes desconocía su condición (proporción de no diagnóstico); esta cifra fue del 65.6% en la franja de 20-39 años y del 18.1% en ≥ 60 años, lo que sugiere rezagos en la detección en la población joven. Otro aspecto que destacar es que entre quienes ya tenían diagnóstico previo, el control (HbA1c < 7%) fue del 36.1% (IC 95%: 26.0-47.6)⁶.

En DT1 pediátrica, los registros del Instituto Mexicano del Seguro Social evidencian fluctuaciones en la incidencia de 2000-2018 (de 3.4 a 2.8 por 100,000; con variaciones por edad y región), lo que destaca la necesidad de estrategias tempranas de detección y manejo⁷.

Para la DG, la adopción de los criterios IADPSG en cohortes urbanas mexicanas eleva la prevalencia estimada a rangos del 10-30% según el punto de corte y la población evaluada, con implicaciones claras para el tamizaje y la carga perinatal^{8,9}.

Este alto porcentaje de pacientes fuera de metas sustenta la necesidad de herramientas que reduzcan la hipoglucemia, optimicen el tiempo en rango (TIR, *time in range*) y faciliten decisiones terapéuticas en tiempo real-objetivos para los que el MCG ha demostrado beneficio consistente en ensayos y en práctica real (GOLD, DIAMOND, HypoDE y MOBILE, entre otros)^{10,13}.

Una proporción importante de los pacientes con DT2 está en tratamiento con insulina, Basu et al. proyectan que alrededor del 7-8% de las personas con DT2 usaría insulina (7.4% en el escenario *status quo*). Con acceso óptimo a la insulina y con

el objetivo de HbA1c < 7, el 15.5% la usaría¹⁴. El reporte de ENSANUT 2012 señala que entre adultos con diabetes que recibían tratamiento farmacológico (85.6%), el 8.0% usaba combinaciones de hipoglucemiantes orales e insulina y el 7.5% usaba solo insulina. Esto equivale a alrededor del 13.3% del total de adultos con diabetes que usan insulina (0.856×0.155). Las proyecciones indican que estas cifras aumentarán considerablemente en los próximos años en nuestro país¹⁵.

RECOMENDACIONES INTERNACIONALES

En general, las guías de práctica clínica (GPC) publicadas en todo el mundo coinciden en las indicaciones para el uso del MCG. Las GPC ADA 2025¹⁶ en su capítulo de tecnologías recomiendan el uso de MCG en tiempo real (rtCGM en inglés) o escaneo intermitente (isCGM en inglés) para todas las personas con DT1 y para personas con DT2 que usan insulina (incluida basal sola, cuando factible), con elección compartida del dispositivo; promueve el uso sistemático de tiempo en rango/tiempo debajo de rango (TIR/TBR) para evaluar y ajustar terapias; sugiere MCG en embarazo con metas de TIR más estrictas, y respalda la integración con sistemas AID (*automated insulin delivery*) en DT1¹⁶.

Las GPC de la Asociación Americana de Endocrinología Clínica 2021 (AACE, *American Association of Clinical Endocrinology*) recomiendan el uso de MCG para DT1 y DT2 insulino dependiente para reducir HbA1c e hipoglucemia, y consideran su uso en DT2 no dependiente de insulina seleccionada para educación y ajuste terapéutico¹⁷. También las GPC de la Sociedad Internacional de Diabetes en niños y adolescentes 2022 (ISPAD, *International Society for Pediatric and Adolescent Diabetes*) recomiendan el empleo de MCG desde el diagnóstico en niños y adolescentes con DT1 cuando sea posible, con objetivos de TIR pediátricos y ajustes por edad/desarrollo^{18,19}.

Las GPC de la Sociedad Mexicana de Nutrición y Endocrinología (SMNE) recomiendan el uso de MCG, especialmente en personas con DT2 que usan insulina, tienen alto riesgo de hipoglucemia, están en ajuste de tratamiento o requieren caracterizar la variabilidad y la TIR, con una recomendación de grado B (evidencia 1-)²⁰. Las GPC publicadas por el Centro Nacional de Excelencia Tecnológica (CENETEC) en 2021, por su parte, no contemplan el uso de tecnologías de MCG en sus recomendaciones²¹.

MÉTODO

Nuestro objetivo fue garantizar el rigor científico, la transparencia y la reproducibilidad, integrando de manera estructurada la mejor evidencia disponible y la experiencia clínica y regulatoria de un grupo multidisciplinario. Este proyecto es una iniciativa de la SMNE, con el objetivo de producir recomendaciones prácticas y alineadas con estándares internacionales, útiles para la toma de decisiones clínicas, la selección de tecnologías y la interpretación de métricas relevantes en contextos reales de atención. La aproximación metodológica se fundamentó en tres pilares complementarios: 1) una búsqueda bibliográfica sistemática en PubMed con estrategias reproducibles basadas en términos MeSH y palabras clave; 2) un proceso de análisis y de extracción de la evidencia científica, y 3) un proceso de consenso mediante un panel Delphi modificado con votación anónima y retroalimentación estructurada utilizando una escala de Likert de 1 a 9, orientado a disminuir sesgos de dominancia y a capturar acuerdo cuantificable entre expertos (Fig. 1).

La gobernanza del proyecto fue designada por la SMNE, responsable de delimitar el alcance del consenso, coordinar el cronograma, supervisar la calidad metodológica y asegurar el cumplimiento de los principios de independencia editorial. El comité directivo actuó como instancia de coordinación y arbitraje metodológico, definiendo procedimientos para el manejo de conflictos de

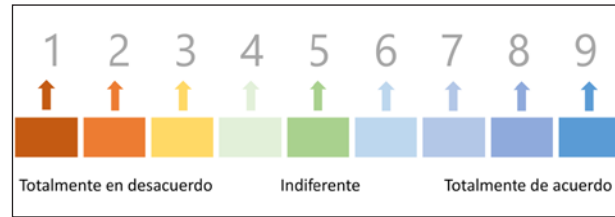


Figura 1. Escala de Likert empleada en Panel Delphi modificado. La escala de Likert evalúa el grado de consenso entre los expertos respecto al contenido y redacción de cada una de las recomendaciones clínicas (*adaptada de Mayorga et al.⁵⁵*).

interés, la validación de estrategias de búsqueda y la conducción de rondas Delphi. Paralelamente, se conformó un grupo de coordinación metodológica con funciones específicas de apoyo técnico, como la construcción de preguntas, el diseño de estrategias de búsqueda, la organización de bases de evidencia, la sistematización de resultados, el análisis de votaciones y la redacción estandarizada.

El panel de expertos multidisciplinario estuvo integrado por especialistas en endocrinología, diabetología, medicina interna y nutrición clínica y, cuando correspondía, en áreas relacionadas con comorbilidades y escenarios especiales. Se buscó incluir perfiles con experiencia clínica en el manejo de personas con diabetes y conocimiento en la interpretación de métricas derivadas de tecnologías de monitoreo. De igual forma, se consideró deseable que el perfil del experto incluyera, además, participación en investigación clínica y familiaridad con los estándares regulatorios y de calidad aplicables a dispositivos médicos. Se consideró como criterio de inclusión la trayectoria profesional relevante (p. ej., años de práctica clínica, experiencia en centros de referencia, participación en comités académicos, publicaciones o contribuciones a guías o consensos, o liderazgo en la implementación de tecnologías). Adicionalmente, se procuró la diversidad geográfica e institucional. Con el fin de fortalecer la independencia y la credibilidad del consenso, se implementó un proceso formal para la declaración y el manejo de conflicto de intereses.

El alcance del consenso se definió mediante la formulación de preguntas estructuradas que

Tabla 1. Documento de alcances

Población incluida
Pacientes con diabetes tipo 1
Pacientes con diabetes tipo 2
Público audiencia
Profesionales de la salud en el primer, segundo y tercer nivel de atención
Elaboradores de políticas públicas
Administrativos y directivos en hospitales públicos y privados
Intervenciones que ser evaluadas
– Tecnologías de dispositivos de monitoreo continuo de glucosa:
– rtCGM (monitoreo en tiempo real)
– isCGM (escaneo intermitente)
Desenlaces que evaluar
– Precisión/utilidad diagnóstica/validez clínica
– Validez/exactitud analítica (<i>point & trend accuracy</i>)
– Desempeño de alertas (rtCGM)
– Fiabilidad/operatividad del sistema
– Reproducibilidad/precisión (<i>between-sensor, within-sensor</i>)
– Seguridad/tolerabilidad
Estándares internacionales
– EE.UU. (FDA)-iCGM (21 CFR 862.1355, Controles especiales)
– CLSI POCT05 (<i>Performance metrics for continuous interstitial glucose monitoring</i>)
– Unión Europea-MDR (CE)

CGM: monitoreo continuo de la glucosa; CLSI: Clinical and Laboratory Standards Institute; FDA: Food and Drug Administration; iCGM: CGM integrable; isCGM: escaneo intermitente; MDR: Medical Device Regulation; rtCGM: monitoreo continuo de la glucosa en tiempo real.

describieran el problema clínico y operativo, el contexto de aplicación y los desenlaces o métricas de interés (Tabla 1). Para garantizar que el consenso respondiera a necesidades reales, las preguntas se construyeron considerando a) escenarios de uso habituales; b) variaciones en la práctica clínica en México; c) requerimientos de comparabilidad en investigación clínica cuando correspondiera, y d) el marco de seguridad y desempeño que subyace a los estándares regulatorios internacionales. Cuando fue posible, las preguntas se formularon siguiendo esquemas tipo PICO (población, intervención, comparador,

desenlaces) o marcos equivalentes adaptados a las tecnologías diagnósticas (Tabla 2). El comité directivo validó el conjunto final de preguntas, priorizando aquellas de mayor impacto en la seguridad, la interpretación clínica y las decisiones terapéuticas.

La búsqueda de evidencia se realizó en PubMed/Medline debido a su cobertura robusta de la literatura biomédica y a su estructuración mediante términos MeSH. Las estrategias de búsqueda se diseñaron de forma iterativa por el grupo metodológico, combinando términos MeSH y texto libre para maximizar la sensibilidad sin sacrificar la especificidad. Se utilizaron operadores booleanos (AND/OR/NOT), truncamientos cuando correspondía y campos específicos (título/resumen, MeSH *terms*) para refinar los resultados. La búsqueda se acompañó de un registro explícito de cada ecuación y de la fecha de ejecución, a fin de permitir su replicación y auditoría.

Las ecuaciones incluyeron términos MeSH y palabras clave relacionadas con la tecnología y su evaluación, por ejemplo: ("Continuous Glucose Monitoring"[MeSH] OR "continuous glucose monitor*" OR CGM OR "intermittently scanned" OR "flash glucose monitoring") AND ("Accuracy"[MeSH] OR accuracy OR precision OR "performance" OR "validation" OR MARD OR "error grid" OR "rate of change" OR "lag") y, cuando correspondió, términos sobre seguridad, usabilidad o aspectos regulatorios. Dependiendo de la pregunta específica, se añadieron filtros por población (p. ej., DT1, DT2, embarazo, hospitalización) o por entorno (ambulatorio, hospitalario). Para minimizar las omisiones, se realizaron búsquedas complementarias a partir de referencias.

Cuando se identificaron revisiones sistemáticas y metaanálisis, se extrajeron medidas de asociación y de efecto (p. ej., diferencias de medias, riesgos relativos, razones de momios (*odds ratios*), intervalos de confianza, medidas de heterogeneidad) y se registró el enfoque de síntesis (modelo de efectos aleatorios o fijos, evaluación del sesgo de publicación cuando estuviera disponible, así como análisis de subgrupos). En el metaanálisis

Tabla 2. Preguntas clínicas

Esquema PICO
¿Qué características deben tener los procedimientos de obtención de datos comparativos (de referencia) para asegurar la validez y la comparabilidad de los resultados en estudios de desempeño de MCG?
¿Cómo debe estructurarse la distribución de las concentraciones de glucosa y de las tasas de cambio en los estudios de desempeño de MCG para reflejar escenarios clínicos relevantes?
¿Qué criterios deben considerarse para la selección y el uso del dispositivo de MCG en estudios clínicos?
¿Cuáles son las métricas recomendadas para evaluar la precisión y la exactitud de los sistemas de MCG en estudios clínicos?
¿Cuál es el papel de la calibración manual capilar en la precisión del MCG?
En dispositivos que requieren calibración manual, ¿cuál es la mejor estrategia de calibración?
¿Cuál es la recomendación, de acuerdo con los estándares internacionales, los consensos y los estándares iCGM de la FDA, respecto a la generación de reportes y el manejo de la información?
¿Cuál es la recomendación, de acuerdo con los estándares internacionales, los consensos y los estándares iCGM de la FDA, respecto a la seguridad del usuario y los servicios de soporte del dispositivo?
Profesionales de la salud en el primer, segundo y tercer nivel de atención
Elaboradores de políticas públicas
Administrativos y directivos en hospitales públicos y privados
Intervenciones que ser evaluadas
– Tecnologías de dispositivos de monitoreo continuo de glucosa: – rtCGM – isCGM
Desenlaces a evaluar
– Precisión/utilidad diagnóstica/validez clínica – Validez/exactitud analítica (<i>point & trend accuracy</i>) – Desempeño de alertas (rtCGM) – Fiabilidad/operatividad del Sistema – Reproducibilidad/precisión (<i>between-sensor, within-sensor</i>) – Seguridad/tolerabilidad – Transferencia y manejo de la información

(Continúa)

Tabla 2. Preguntas clínicas (continuación)

Estándares Internacionales
– EE.UU. (FDA)-iCGM (21 CFR 862.1355, Controles especiales) – CLSI POCT05 (<i>Performance metrics for continuous interstitial glucose monitoring</i>) – Unión Europea-MDR (CE)

CGM: monitoreo continuo de la glucosa; CLSI: Clinical and Laboratory Standards Institute; FDA: Food and Drug Administration; iCGM: CGM integrable; isCGM: escaneo intermitente; MDR: *medical device regulation*; rtCGM: monitoreo continuo de la glucosa en tiempo real.

en red se registraron elementos adicionales, como comparaciones indirectas, consistencia/transitividad y *ranking* (cuando se reportó), lo que enfatiza la interpretación clínica de los estimadores y su aplicabilidad en el contexto.

El proceso de síntesis se estructuró para facilitar su integración en el Delphi modificado. El grupo metodológico elaboró un borrador en lenguaje técnico y clínicamente operativo, acompañado de un breve racional y de las referencias correspondientes. La redacción de enunciados se diseñó para ser evaluable mediante una escala ordinal de acuerdo, evitando la ambigüedad excesiva y, cuando fue necesario, separando conceptos complejos en enunciados independientes. Cuando la evidencia era consistente y robusta, los enunciados se formularon como recomendaciones afirmativas; cuando existía incertidumbre o heterogeneidad sustancial, se emplearon formulaciones más condicionadas, preservando precisión y evitando extrapolaciones no justificadas.

El consenso formal se condujo mediante un Delphi modificado, ya que combinó rondas de votación anónima con espacios estructurados de retroalimentación y discusión técnica, sin que la discusión sustituyera la votación individual; es decir, la deliberación se utilizó para clarificar las evidencias y resolver las interpretaciones, pero la decisión final se basó en votaciones anónimas y repetidas, con retroalimentación agregada. La evaluación de cada enunciado se realizó mediante una escala de Likert de 1 a 9, donde los valores bajos reflejaron desacuerdo y los valores

altos, acuerdo, siguiendo un esquema ordinal ampliamente utilizado en procesos de consenso formales. De manera operativa, los valores 1-3 se interpretaron como desacuerdo (o «no recomendado/inapropiado»), los valores 4-6 como incertidumbre o acuerdo parcial, y los valores 7-9 como acuerdo (o «recomendado/apropiado»). Para cada enunciado, además de la calificación numérica, los participantes pudieron registrar comentarios cualitativos estructurados: justificación de su voto, sugerencias de redacción, solicitudes de evidencia adicional, identificación de supuestos implícitos o de condiciones de aplicabilidad. Los expertos calificaron de forma independiente y anónima el conjunto inicial de enunciados. La anonimización se mantuvo para minimizar la influencia interpersonal y promover votaciones auténticas. Tras la ronda, el equipo metodológico analizó los resultados, calculando estadísticos descriptivos para cada enunciado (mediana, rango intercuartílico, distribución de frecuencias por categorías 1-3/4-6/7-9) e identificando patrones de desacuerdo. En consonancia con las aproximaciones Delphi formales, se definieron criterios de consenso que combinaron la centralidad y la dispersión.

La retroalimentación para la segunda ronda, en los casos en que no se alcanzó un porcentaje mínimo de consenso (70%), se construyó de manera sistemática. Para cada enunciado, se compartió con el panel: a) la mediana y la distribución de votos (por categorías); b) un resumen temático de los comentarios (anonimizado), agrupando argumentos a favor y en contra; c) aclaraciones o evidencia adicional si se identificó la necesidad (p. ej., la incorporación de datos cuantitativos extraídos de revisiones sistemáticas o metaanálisis), y d) una versión revisada del enunciado cuando la redacción podía mejorar la precisión o la condicionalidad.

En consonancia con las buenas prácticas de los consensos contemporáneos, se consideró la importancia de la actualización. Dado que el campo de tecnologías de monitoreo y de estándares de desempeño evoluciona con rapidez, se recomendó que el consenso incluyera un plan de revisión periódica (p. ej., cada 24-48 meses).

CRITERIOS MÍNIMOS DE PRECISIÓN DIAGNÓSTICA PARA SISTEMAS DE MONITOREO CONTINUO DE GLUCOSA

El consenso mexicano propone establecer un conjunto de criterios mínimos de precisión diagnóstica que permitan evaluar de manera uniforme, transparente y clínicamente relevante el desempeño de los sistemas de MCG en el país. El objetivo es asegurar que cualquier tecnología que busque incorporarse al mercado mexicano cumpla con estándares que garanticen la seguridad, la confiabilidad y la aplicabilidad en condiciones reales de uso (Tabla 3).

Requerimientos operacionales de precisión en condiciones ambulatorias

El consenso recomienda que la evaluación de la precisión de los sistemas de MCG se extienda más allá de los criterios analíticos tradicionales y se complemente con métricas operacionales obtenidas en condiciones ambulatorias reales. Para lograrlo, se establece que todo sistema de MCG deberá reportar, de forma obligatoria:

Indicadores derivados del comportamiento glucémico cotidiano, calculados sobre un mínimo de 14 días con $\geq 70\%$ de las lecturas esperadas, incluyendo:

- TIR (*time in range*, tiempo en rango): 70-180 mg/dl
- TBR (*time below range*, tiempo abajo de rango): < 70 mg/dl y < 54 mg/dl, reportados por separado
- TAR (*time above range*, tiempo arriba de rango): > 180 mg/dl y > 250 mg/dl, reportados por separado
- Glucosa media del sensor
- Variabilidad glucémica: coeficiente de variación (CV), con referencia de estabilidad $\leq 36\%$
- GMI (*glucose management indicator*, indicador de manejo de glucosa)

Tabla 3. Estándares y criterios específicos para CGM

EE.UU. (FDA)-iCGM (21 CFR 862.1355, Controles especiales): define desempeño/seguridad, validación clínica, interoperabilidad, ciberseguridad, etiquetado (remite a 21 CFR 809.10), gestión de riesgo, usabilidad y requisitos de ciclo de vida. Es la referencia de calidad más concreta para CGM «integrables» (iCGM)
CLSI POCT05 (Performance metrics for continuous interstitial glucose monitoring): métrica y métodos para evaluar exactitud (p. ej., MARD), sesgo vs. referencia, latencia y desempeño por rangos de glucosa/ROC. Reconocida por FDA como consenso voluntario
ISO/IEEE 11073-10425 (perfil para intercambio de datos CGM con sistemas personales): especifica interoperabilidad de datos/telemetría entre sensor-transmisor-receptores
IEC 60747-14-10:2019 (<i>Semiconductor devices-Wearable electrochemical-glucose sensors</i>): métodos de evaluación de desempeño a nivel de sensor electroquímico portátil
prEN/IEC 60601-2-95 (en desarrollo): estándar particular de seguridad/desempeño para CGM dentro de la familia 60601; útil para planificar evidencia técnica futura en CE/UKCA

Nota: fuera de FDA/CLSI y las IEC/IEEE anteriores, no existe aún un ISO «particular» de desempeño clínico para CGM equivalente a ISO 15197 (que es para glucometría capilar, no CGM). El aseguramiento de calidad de CGM en otras regiones se demuestra con normas transversales (riesgo, *software*, biocompatibilidad, seguridad eléctrica, EMC, usabilidad, QMS) + evidencia clínica acorde al marco local.

CGM: monitoreo continuo de la glucosa; CLSI: Clinical and Laboratory Standards Institute; EMC: compatibilidad electromagnética; FDA: Food and Drug Administration; iCGM: CGM integrable, *integrated continuous glucose monitoring*; IEC: International Electrotechnical Commission; IEEE: Institute of Electrical and Electronics Engineers; isCGM: escaneo intermitente; ISO: International Organization for Standardization; MARD: *mean absolute relative difference*, diferencia relativa absoluta media; MDR: *medical device regulation*; QMS: sistema de gestión de calidad; rtCGM: monitoreo continuo de la glucosa en tiempo real.

Modificado de ElSayed et al.¹⁶, Klonoff et al.³⁷ y Scharf et al.³⁹

– Porcentaje de tiempo con sensor activo (captura de datos)

Estas métricas reflejan directamente la utilidad clínica del sistema y permiten evaluar si el desempeño reportado en estudios controlados se mantiene en escenarios reales de uso. La incorporación de estos indicadores como requerimientos mínimos responde a la necesidad de evaluar no solo la exactitud analítica del sensor, sino también su capacidad para apoyar decisiones terapéuticas en la vida diaria. El consenso reconoce que estudios contemporáneos ya utilizan de manera consistente estas métricas como indicadores centrales de desempeño en la vida real, lo que demuestra su factibilidad y relevancia clínica. Sobre esta base, se propone que México adopte estas métricas como parte esencial del estándar mínimo para la evaluación operativa de nuevos sistemas de MCG.

Evaluación obligatoria del desempeño en extremos glucémicos

Con el fin de garantizar la seguridad de las personas usuarias de tecnologías de MCG, el consenso establece que los sistemas de MCG deberán demostrar

precisión diferenciada en rangos glucémicos extremos. Esto incluye, de manera obligatoria:

- Desempeño en hipoglucemia (< 70 mg/dl), con énfasis en la caracterización de niveles 1 y 2 cuando la evidencia disponible lo permita.
- Desempeño en hiperglucemia marcada (> 250 mg/dl) y, cuando sea posible, en valores superiores a 300 mg/dl.

La exigencia de precisión en estos rangos responde a la realidad clínica de México, donde una proporción importante de personas con diabetes experimenta variabilidad glucémica significativa, episodios de hipoglucemia y periodos prolongados de hiperglucemia. Los errores de medición en estas zonas críticas pueden tener un impacto directo en la toma de decisiones terapéuticas, especialmente en personas que utilizan esquemas intensivos de insulina, bombas de infusión o sistemas automatizados.

Al incluir estos requerimientos, el consenso fortalece el marco de evaluación de los MCG al asegurar que la exactitud reportada no se limite a valores cercanos a la euglucemia, sino que abarque los escenarios en los que la precisión tiene mayor impacto clínico.

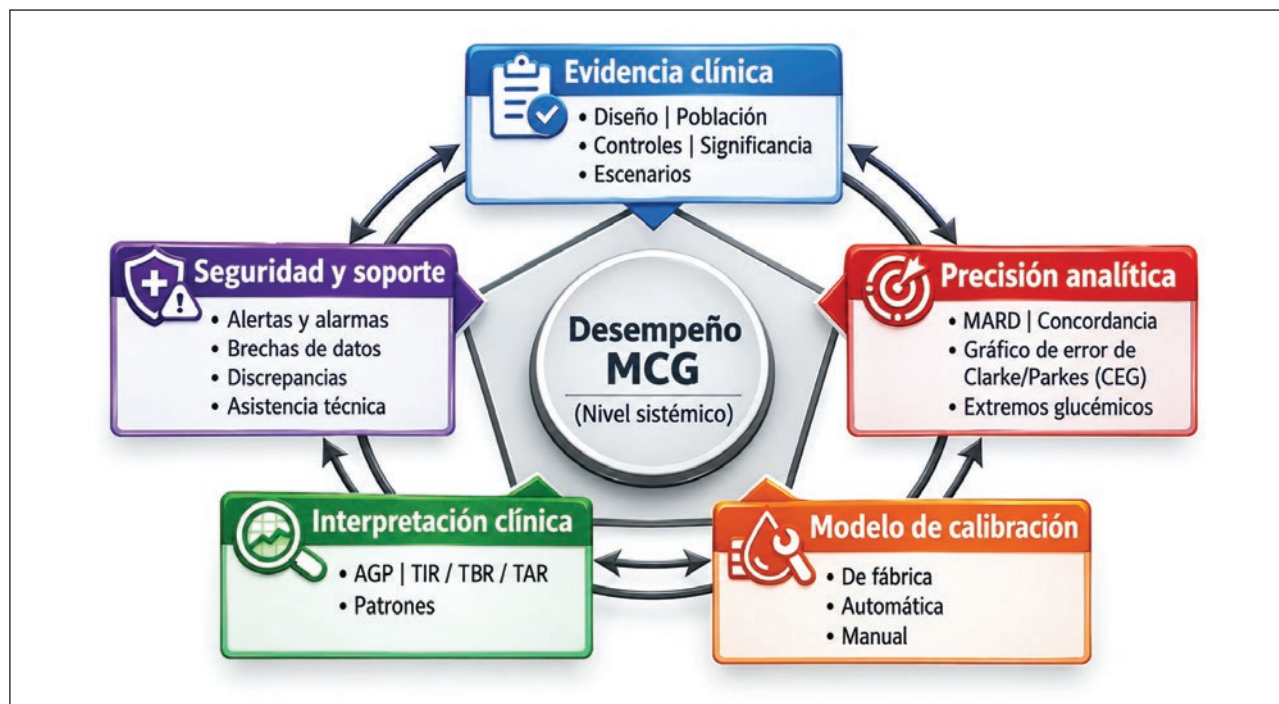


Figura 2. Desempeño MCG para evaluación de sistemas de CGM/MCG. AGP: *ambulatory glucose profile*, perfil ambulatorio de glucosa; CGM: *continuous glucose monitoring*, monitoreo continuo de glucosa (MCG); MARD: *mean absolute relative difference*, diferencia relativa absoluta media; TAR: *time above range*, tiempo arriba de rango; TBR: *time below range*, tiempo abajo de rango; TIR: *time in range*, tiempo en rango.

INTEGRACIÓN DE MÉTRICAS OPERACIONALES Y EXTREMAS EN UN ESTÁNDAR NACIONAL MÍNIMO

El consenso propone que los criterios descritos (métricas ambulatorias de desempeño y de precisión en los extremos glucémicos) formen parte de un estándar nacional mínimo para la evaluación de nuevas tecnologías de MCG en México. Esta integración permitiría:

- Alinear la evaluación de desempeño con prácticas clínicas actuales.
- Facilitar comparaciones entre dispositivos.
- Garantizar que la decisión de incorporar nuevas tecnologías se base en información relevante para la vida real.
- Ofrecer a reguladores, clínicos y sistemas de salud un marco uniforme para evaluar seguridad y utilidad (Fig. 2).

Este enfoque no solo fortalece la protección de las personas usuarias, sino que también posiciona a México en un proceso regulatorio más avanzado, coherente con las tendencias internacionales que buscan armonizar la precisión analítica con el desempeño clínico real (Tabla 4).

PREGUNTA 1. ¿QUÉ CARACTERÍSTICAS DEBEN TENER LOS PROCEDIMIENTOS DE OBTENCIÓN DE DATOS COMPARATIVOS (DE REFERENCIA) PARA ASEGURAR LA VALIDEZ Y LA COMPARABILIDAD DE LOS RESULTADOS EN ESTUDIOS DE DESEMPEÑO DE MONITOREO CONTINUO DE GLUCOSA?

Texto de soporte y análisis

La validez y comparabilidad de los estudios de desempeño de los sistemas de MCG dependen

Tabla 4. Matriz regional de estándares para CGM

Jurisdicción	Marco regulatorio	Estándar/guía específica CGM ¹	Exactitud/desempeño
EE.UU. (FDA)	21 CFR (862, 809), QSR 21 CFR 820/QMS	iCGM (21 CFR 862.1355 - Controles especiales)	CLSI POCT05; validación clínica iCGM
Unión Europea (MDR)	MDR 2017/745 (CE)	IEC 60601-2-95 (en desarrollo)	Demostración clínica conforme GSPR
Reino Unido (MHRA/UKCA)	UK MDR 2002 (enmiendas), Designated Standards	S/E (no particular; seguir paquete transversal)	CLSI POCT05 (acceptable)
Canadá (Health Canada)	CMDR; Recognized Standards; MDSAP	S/E (no particular; seguir paquete transversal)	CLSI POCT05 (acceptable)
Australia (TGA)	Essential Principles + Standards Orders	S/E (no particular; seguir paquete transversal)	CLSI POCT05 (acceptable)
Japón (PMDA/MHLW)	Pharmaceuticals and Medical Devices Act; JIS/IEC reconocidos	S/E (no particular; seguir paquete transversal)	CLSI POCT05 (acceptable)
Brasil (ANVISA/INMETRO)	RDC 751/2022; Portaria INMETRO 384/2020	S/E (no particular; seguir paquete transversal)	CLSI POCT05 (acceptable)
México (COFEPRIS)	NOM-241-SSA1-2021 (BPF/QMS)	S/E (no particular; seguir paquete transversal)	CLSI POCT05 (acceptable)

ANVISA: Agencia Nacional de Vigilancia Sanitaria; CGM: monitoreo continuo de la glucosa; CLSI: Clinical and Laboratory Standards Institute; COFEPRIS: Comisión Federal para la Protección contra Riesgos Sanitarios; EMC: *electromagnetic compatibility*; FDA: Food and Drug Administration; iCGM: CGM integrable, *integrated continuous glucose monitoring*; IEC: International Electrotechnical Commission; IEEE: Institute of Electrical and Electronics Engineers; INMETRO: Instituto Nacional de Metrología, Calidad y Tecnología; isCGM: escaneo intermitente; ISO: International Organization for Standardization; MARD: *mean absolute relative difference*, diferencia relativa absoluta media; MDR: *medical device regulation*; MHLW: Ministry of Health, Labour and Welfare; MHRA: Medicines and Healthcare products Regulatory Agency; PMDA: Pharmaceuticals and Medical Devices Agency; QMS: *Quality Management System*; rtCGM: monitoreo continuo de la glucosa en tiempo real; TGA: Therapeutic Goods Administration; UKCA: UK Conformity Assessed.

Adaptada de Scharf M et al.³⁹

de manera crítica de cómo se obtienen, miden y gobiernan los datos del comparador (referencia), porque el desempeño observado del MCG puede cambiar de forma sustancial al modificar el origen de la muestra, el método analítico, la frecuencia de muestreo, la alineación temporal y la trazabilidad metrológica del procedimiento comparativo, incluso cuando el MCG evaluado es el mismo. Esta dependencia metodológica se reconoce explícitamente como un obstáculo histórico para la estandarización: la ausencia de procedimientos comparativos armonizados favorece resultados difícilmente comparables entre estudios y confunde la interpretación clínica de métricas como la MARD (*mean absolute relative difference*, diferencia relativa absoluta media), una métrica utilizada para cuantificar la exactitud numérica de un

sistema de MCG al compararlo con una medición de referencia²².

Un principio rector indispensable es aceptar que el comparador no está «libre de error»: toda medición comparativa está afectada por sesgos e imprecisión, y no es posible extraer de manera perfecta el error del comparador del error del MCG; en consecuencia, cualquier variación sistemática del comparador entre estudios (o dentro de un mismo estudio) se traduce en una variación artificial del desempeño del MCG. Por ello, los procedimientos de obtención de datos comparativos deben definirse como un sistema integral (preanalítico-analítico-postanalítico) con especificaciones *a priori*, controles de desempeño y reglas de corrección, de modo que el comparador se aproxime metrológicamente a un «valor verdadero» operativo y, sobre

todo, que lo haga de manera reproducible entre centros, periodos y dispositivos²³.

La primera característica que debe estandarizarse es el origen de la muestra (capilar, venosa o venosa «arterializada»), porque estos compartimentos son fisiológicamente distintos y esa diferencia se amplifica ante cambios rápidos en la glucosa. En evidencia reciente, la diferencia promedio entre glucosa capilar y venosa se ha descrito en el rango del 5-10% y puede variar con la velocidad de cambio; además, la arterialización (calentamiento del brazo/mano) puede modificar la magnitud de esa diferencia (p. ej., una diferencia capilar-venosa de +5.9% reducida a +4.2% con calor), con repercusión directa sobre el sesgo observado y la MARD del MCG²³.

La selección del origen de la muestra debe alinearse con la pregunta regulatoria y de uso clínico que se pretende sustentar: se ha propuesto que para el desempeño en escenarios de uso real del MCG, el comparador capilar optimiza la alineación con la verificación clínica mediante autocontrol (SMBG, *self-monitoring of blood glucose*), mientras que la glucosa venosa es el ancla típica de diagnóstico y de múltiples decisiones en entornos sanitarios²³.

Un segundo componente crítico es la matriz (sangre total, plasma o suero) y la gestión preanalítica. Aunque la glucosa en sangre total puede convertirse en valores equivalentes a los del plasma, esa conversión es una aproximación que puede verse influida por factores como el hematocrito y la enfermedad. De manera particularmente importante, la medición en suero se considera inadecuada en este contexto debido al impacto de la glucólisis (pérdida de glucosa durante el metabolismo celular hasta el momento del análisis), lo que puede sesgar sistemáticamente la referencia. En términos operativos, esto obliga a procedimientos comparativos que especifiquen: tipo de tubo/anticoagulante, tiempo máximo desde la extracción hasta la lectura, condiciones de transporte/temperatura, reglas de descarte por demora y trazabilidad de conversiones (si se usan), además de documentar covariables

biológicas relevantes (p. ej., hematocrito cuando se infiere conversión)²³.

En paralelo, el procedimiento comparativo debe definir y auditar la jerarquía del método analítico y su trazabilidad metrológica. Se ha propuesto formalizar especificaciones analíticas de desempeño para comparadores (C-APS, por sus siglas en inglés), compuestas por sesgo e imprecisión, definibles con base en la variación biológica. Se han recomendado criterios de sesgo $\leq 2.1\%$ e imprecisión $\leq 2.4\%$ a lo largo del rango de medición, junto con trazabilidad hacia métodos/materiales de orden metrológico superior. Estos requerimientos no son triviales para ciertos dispositivos SMBG, lo que obliga a que el procedimiento comparativo sea explícito respecto al tipo de instrumento (analizador de laboratorio, SMBG), a su validación interna contra C-APS y a sus limitaciones por interferencias²³.

Para que el control de sesgo no sea puramente teórico, el procedimiento comparativo debe estandarizar las réplicas y la estimación de la imprecisión en las muestras de participantes. Se recomienda medir las muestras comparadoras al menos por duplicado y usar el promedio como valor de referencia, tanto para reducir la imprecisión como para detectar valores atípicos (*outliers*) en la muestra clínica²³.

La frecuencia de muestreo comparativo y la cobertura del espacio glucémico constituyen otra dimensión esencial de la validez. Se ha recomendado un esquema «típico» de muestreo cada 15 minutos durante 6-8 horas (aprox. 25-33 extracciones consecutivas) en sesiones controladas, y no como un evento único, sino distribuido estratégicamente a lo largo de la vida útil del sensor (p. ej., inicio, mitad y final), para captar cambios de desempeño por calentamiento, deriva o envejecimiento²³.

En una revisión sistemática (RS) centrada en la precisión y la confiabilidad de MCG en entornos ambulatorios, los autores extrajeron explícitamente características del estudio de referencia, incluyendo la técnica de laboratorio y el intervalo temporal entre pruebas, precisamente porque

estas variables determinan la credibilidad de las métricas de exactitud. Se reportó una exactitud promedio global (como MARD) del 9.4%, con sensibilidad y especificidad promedio para hipoglucemia del 85.7 y 95.33%, y para hiperglucemia del 97.45 y 96%, respectivamente; sin embargo, el mismo trabajo concluye que la exactitud es adecuada en euglucemia/hiperglucemia, pero permanece como reto en hipoglucemia, lo cual tiene implicaciones directas para el comparador: si el muestreo comparativo es escaso o de baja calidad en el rango bajo, se corre el riesgo de sobreestimar desempeño o de producir estimaciones imprecisas y heterogéneas entre estudios²⁴.

La necesidad de adaptar el comparador al contexto clínico es particularmente evidente en escenarios perioperatorios e intrahospitalarios, donde el «comparador natural» suele ser la medición capilar, pero su fiabilidad puede degradarse por condiciones hemodinámicas, perfusión periférica o interferencias farmacológicas; a la vez, el propio MCG puede verse afectado por edema, compresión del sitio o hipoperfusión grave. En una RS sobre MCG perioperatorio en cirugía no cardíaca, se destaca que la vigilancia continua del funcionamiento del sensor es crucial cuando los cambios agudos comprometen la confiabilidad del dispositivo. Este tipo de evidencia obliga a que el procedimiento de obtención de datos comparativos (referencia) incluya: la selección de un método comparador menos vulnerable a artefactos del entorno, la documentación de cofactores clínicos que modulan el sesgo (perfusión, edema, fármacos), y reglas predefinidas para pausar, repetir o recalibrar el procedimiento comparativo cuando existan condiciones que invaliden la medición de referencia²⁵.

Otra RS específicamente orientada al uso y la exactitud del MCG en el periodo perioperatorio sintetiza que, pese al crecimiento del interés, persisten incertidumbres metodológicas y la necesidad de evidencia robusta sobre efectividad, seguridad y precisión en ese entorno. Aunque la pregunta clínica primaria de ese cuerpo de evidencia puede no ser metrológica, su implicación para los procedimientos

comparativos es directa: en contextos de variaciones rápidas, perfusión cambiante y decisiones críticas, el comparador debe tener un desempeño analítico y una logística de muestreo compatibles con la toma de decisiones, y no debe degradarse por conveniencia operativa²⁶.

La comparabilidad también depende de reconocer que el MCG produce series temporales de alta granularidad (p. ej., cada 5 minutos), lo que plantea un requisito implícito para el comparador: capturar, con suficiente densidad, el comportamiento dinámico de la glucosa y permitir métodos de pareamiento temporal que no destruyan información. En una RS sobre técnicas estadísticas para analizar MCG simultáneo con electrocardiografía ambulatoria se enfatiza que la agregación excesiva de series densas implica pérdida de potencia y de información; trasladado a estudios de desempeño, esto significa que un comparador medido esporádicamente puede subestimar o sobreestimar errores si no intercepta adecuadamente excursiones, pendientes y extremos. Así, el procedimiento comparativo debe: a) mantener una frecuencia que permita evaluar el desempeño no solo en «estado estable» sino también durante transiciones; b) registrar con precisión las marcas de tiempo para el pareamiento, y c) preespecificar si se emplearán enfoques basados en valores puntuales pareados, ventanas o métricas dependientes de la tasa de cambio²⁷.

En el mismo sentido, la heterogeneidad de herramientas analíticas para MCG documentada en una RS de paquetes y algoritmos estadísticos subraya que la comparabilidad no se logra solo midiendo más veces, sino estandarizando qué se hace con los datos faltantes, cómo se imputan (si se imputan), qué reglas de exclusión existen y cómo se reportan las métricas. Dado que un porcentaje relevante de paquetes puede imputar datos faltantes, existe el riesgo de que diferencias en el manejo de los datos faltantes conviertan la calidad del comparador en un problema invisible: si la referencia o el MCG tienen huecos y cada estudio los trata de forma distinta, el desempeño resultante deja de ser comparable²⁸.

La importancia del comparador se aprecia con especial claridad cuando el MCG se contrasta con estándares clínicos «de diagnóstico» (no solo con la glucosa sanguínea instantánea). En una RS que comparó MCG con la prueba de tolerancia oral a la glucosa (OGTT, *oral glucose tolerance test*) para detectar disglucemia/diabetes en fibrosis quística, el resultado principal mostró que la diabetes diagnosticada por MCG tuvo un riesgo relativo (RR) de 2.92 (IC 95%: 2.20-3.88) frente a la diagnosticada por OGTT²⁹.

De manera análoga, en una RS y metaanálisis de exactitud diagnóstica de dispositivos en el sitio de cuidado para hipoglucemia neonatal, un contexto donde el error de referencia tiene consecuencias directas, se documenta que sensibilidad y especificidad varían sustancialmente según el principio analítico (p. ej., glucosa oxidasa vs. deshidrogenasa; fotometría vs. electroquímica), con ejemplos de estimaciones agrupadas como sensibilidad 0.73 y especificidad 0.99 para métodos enzimáticos con electroquímica, o sensibilidad 0.92 y especificidad 0.92 para métodos no enzimáticos con fotometría, entre otras combinaciones. Este estudio señala que la elección del método del comparador (y su control de sesgo/imprecisión) no es neutra; condiciona la capacidad para clasificar correctamente eventos (sobre todo en rangos bajos) y puede introducir heterogeneidad sistemática entre estudios³⁰.

Una RS sobre tecnologías no invasivas de monitoreo de glucosa refuerza un punto de fundamento metrológico: el método de referencia para la concentración de glucosa se ha descrito como el método de hexocinasa, pero su implementación típicamente requiere equipamiento de laboratorio y volúmenes que no siempre son compatibles con diseños de muestreo frecuente, lo que impulsa el uso de alternativas (capilares) y aumenta la probabilidad de heterogeneidad comparativa entre estudios. Desde la perspectiva de validez y comparabilidad en desempeño de MCG, esta tensión obliga a explicitar en el procedimiento comparativo: a) por qué se selecciona un comparador factible (p. ej., capilar) sin renunciar a trazabilidad (p. ej., corrección retrospectiva

contra orden metrológico superior); b) cómo se controlará el impacto de matriz/origen sobre el sesgo, y c) cómo se garantizará que la referencia no sea el «eslabón débil» que limite la inferencia sobre el MCG³¹.

En síntesis, los procedimientos de obtención de datos comparativos para estudios de desempeño de MCG deben incorporar, como mínimo, la estandarización explícita del origen de la muestra (sin mezclar capilar y venosa salvo diseño preespecificado), el control preanalítico y la elección de una matriz adecuada, evitando suero por glucólisis, así como la selección de métodos comparadores con trazabilidad metrológica y el cumplimiento de especificaciones analíticas. La medición por duplicado con reglas de valores atípicos (*outliers*), auditoría del sesgo del comparador y corrección retrospectiva cuando proceda, frecuencia de muestreo suficiente y distribución garantizada de valores en rangos extremos y en alta tasa de cambio (RoC, *rate of change*), además de alineación temporal estricta y reporte transparente de flujo y tiempos, y manejo de datos faltantes para preservar la comparabilidad; también son aspectos muy relevantes^{22,23}.

PREGUNTA 2. ¿CÓMO DEBE ESTRUCTURARSE LA DISTRIBUCIÓN DE LAS CONCENTRACIONES DE GLUCOSA Y DE LAS TASAS DE CAMBIO EN LOS ESTUDIOS DE DESEMPEÑO DE MONITOREO CONTINUO DE GLUCOSA PARA REFLEJAR ESCENARIOS CLÍNICOS RELEVANTES?

Texto de soporte y análisis

La validez y la comparabilidad entre estudios de desempeño de sistemas de MCG dependen de que la «exposición» del sistema por evaluar sea homogénea, es decir, que los pares MCG-referencia se recolecten bajo una distribución predefinida y clínicamente relevante de concentraciones de glucosa y de RoC. Esto es crítico porque la

exactitud del MCG no es constante a lo largo del rango glucémico ni bajo todas las dinámicas: las condiciones de hipoglucemia, hiperglucemia marcada y transiciones rápidas (ascensos/descensos) son escenarios en los que se concentran decisiones terapéuticas de alto impacto (p. ej., correcciones, rescates, ajuste de insulina basal/bolo, suspensión), y también en los que se ha observado potencial deterioro de desempeño si están subrepresentados en el muestreo comparativo. En consecuencia, si la distribución de datos comparativos se concentra en exceso en regiones «neutras» (euglucemia relativamente estable) o si la RoC no se caracteriza ni se estandariza, los estimadores de desempeño (p. ej., MARD, sesgo, métricas de alerta) pueden verse sesgados^{23,32}.

La necesidad de anclar la distribución de glucosa y RoC a escenarios clínicamente relevantes se entiende mejor cuando se observa que las métricas derivadas de MCG que hoy guían los objetivos terapéuticos y la evaluación de intervenciones son sensibles a errores diferenciales por región glucémica y por dinámica. En una RS con metaanálisis y metaanálisis en red en adultos con DT1, la tecnología MCG se asoció con una reducción de HbA1c (diferencia de medias [DM]: -0.38% ; IC 95%: -0.49 a -0.27) y un aumento de TIR (DM: 7.9% ; IC 95%: 5.8 - 10.0), equivalente a 114 minutos/días adicionales en rango. Además, el MCG en tiempo real mostró ventajas frente a MCG de escaneo intermitente para TIR ($+5.63\%$) y la reducción de la hipoglucemia (-1.28%), y los sistemas automatizados obtuvieron la mejor jerarquización (SUCRA: 0.92), con mayor beneficio en quienes tenían HbA1c basal más elevada (DM $> 8.5\%$: -0.68 vs. $< 7.5\%$: -0.24 ; $p = 0.009$). Este patrón de beneficios se expresa precisamente en «zonas frontera» (alrededor de umbrales y durante cambios)³³.

De forma consistente, una RS con metaanálisis de ensayos controlados aleatorizados (ECA) que comparó rtCGM vs. isCGM en población con DT1 encontró que la diferencia entre tecnologías se expresó sobre todo en métricas de «tiempo en estados» más que en HbA1c: el MCG en tiempo real incrementó el TIR en 7.0% (IC 95%: 5.8 - 8.3) y

redujo el TBR < 70 mg/dl en -1.7% (IC 95%: -3.0 a -0.4), sin diferencias detectables en HbA1c³⁴.

La relevancia de incorporar RoC de manera explícita se acentúa al considerar la evidencia sintetizada sobre las tecnologías AID. En una RS y metaanálisis de ECA ambulatorios en niños y adolescentes con DT1, los sistemas AID se asociaron con un incremento del TIR de 11.38% (IC 95%: 9.01 - 13.76), una reducción del TBR del -0.59% (IC 95%: -1.02 a -0.15) y una reducción del TAR del -12.19% (IC 95%: -14.65 a -9.73), con efectos sostenidos a los 3 y 6 meses. Dado que los algoritmos AID dependen de señales de tendencia y de anticipación (que conceptualmente se relacionan con RoC y con la proximidad a umbrales), una evaluación de desempeño de MCG que aspire a ser clínicamente representativa debe ir más allá de «cuántos puntos» están en hipoglucemia o hiperglucemia: debe asegurar también que existan suficientes pares de datos en condiciones de ascenso/descenso rápidos³⁵.

La justificación empírica para enriquecer deliberadamente las regiones de hipoglucemia y de cambios rápidos también proviene de la literatura sobre exactitud y fiabilidad. En una RS sobre precisión y fiabilidad de MCG en tiempo real en entornos ambulatorios (22 estudios; 2,294 pacientes), se reportó una MARD promedio del 9.4% , pero se observó que, al aplicar como referencia comparativa los criterios de exactitud de la norma ISO 15197:2013 (diseñada para glucómetros capilares, no para MCG), ningún dispositivo evaluado los alcanzó en el dominio de hipoglucemia ($\geq 95\%$ dentro de ± 15 mg/dl), mientras que dos sí lo lograron en hiperglucemia (Dexcom G6 y G7). El criterio regulatorio aplicable a MCG en hipoglucemia es el de iCGM ($\geq 85\%$ dentro de ± 15 mg/dl). Adicionalmente, la capacidad de detección de eventos (que depende críticamente del desempeño cerca de umbrales) mostró, para hipoglucemia, una sensibilidad del 85.7% y especificidad del 96.1% (valor predictivo positivo [VPP] 73.6% , valor predictivo negativo [VPN] 98.4%), y, para hiperglucemia, una sensibilidad del 91.1% y especificidad del 92.3% (VPP 80.0% , VPN 97.3%)²⁴.

En una revisión del alcance de los estudios de desempeño clínico (129 estudios) se documentó

que, incluso cuando se reporta la distribución por rangos, la estandarización es limitada: entre los estudios que informaron valores comparativos del tipo TIR, la mediana de los valores del comparador fue $7.3\% < 70 \text{ mg/dl}$, 61.4% entre $70\text{-}180 \text{ mg/dl}$ y $31.4\% > 180 \text{ mg/dl}$, con variabilidad considerable. Más crítico aún para la pregunta de RoC, la revisión subraya que el cálculo de RoC solo es significativo si la frecuencia de muestreo es suficientemente alta; usando un umbral operativo de 20 minutos o menos se identificó que de 83 estudios con intervalos de muestreo $\leq 20 \text{ min}$, solo 21 (25.3%) reportaron una distribución de RoC del comparador. Para superar esta inconsistencia, se propuso reportar la tasa absoluta media de cambio (MARoC, por sus siglas en inglés) como un resumen numérico; la mayoría de los valores de MARoC se ubicaron entre $0.7\text{-}1.0 \text{ mg/dl/min}$, con algunos estudios por encima de 1.3 mg/dl/min ²³.

En respuesta a esta brecha, se ha propuesto conceptual y operativamente que la distribución de datos comparativos no debe definirse únicamente por rangos de concentración (p. ej., < 70 , $70\text{-}180$, > 180), sino por la combinación concentración-RoC, porque la interpretación clínica depende del contexto dinámico: el mismo descenso (p. ej., -3 mg/dl/min) no tiene la misma implicación clínica si ocurre a 90 mg/dl que si ocurre a 230 mg/dl , y un valor de 72 mg/dl exige decisiones distintas si la tendencia es ascendente ($+2 \text{ mg/dl/min}$) o descendente (-2 mg/dl/min). A partir de esta premisa, se desarrolló una representación gráfica de la combinación «glucosa en sangre (BG)-RoC», denominada región dinámica de glucosa (DGR, por sus siglas en inglés), que clasifica pares BG-RoC en regiones críticas vs. neutras, buscando alinear la evaluación del MCG con los escenarios donde la exactitud y la confiabilidad de las alertas son más determinantes³².

En términos de concentración aislada, la recomendación explícita para armonizar los estudios de desempeño es que, durante la evaluación, al menos el 7.5% de las concentraciones de BG del comparador sean $< 70 \text{ mg/dl}$ (3.9 mmol/l) y al menos 7.5% sean $> 300 \text{ mg/dl}$ (16.7 mmol/l), respectivamente.

Este umbral cumple dos objetivos: primero, obliga a que el conjunto de datos no esté dominado por euglucemia relativamente estable (donde muchos sistemas lucen bien); y segundo, asegura densidad suficiente en los extremos donde los errores tienen consecuencias clínicas desproporcionadas (hipoglucemia) o donde la calibración/linealidad y los retardos fisiológicos pueden afectar el acuerdo (hiperglucemia marcada)³².

Para operacionalizar la dimensión dinámica, la propuesta define regiones críticas en el plano BG-RoC con porcentajes mínimos e introduce explícitamente el concepto de «hipoglucemia/hiperglucemia inminente». En la tabla de definición de regiones del DGR, se recomienda que: 1) la región de glucosa baja $< 70 \text{ mg/dl}$, cualquier RoC represente $\geq 7.5\%$ de los pares; 2) la región de glucosa alta $> 300 \text{ mg/dl}$, cualquier RoC represente $\geq 7.5\%$; 3) la región de alerta baja (*alert low*) represente $\geq 7.5\%$ y se define como $\text{BG} \geq 70 \text{ mg/dl}$ con $\text{RoC} < -1 \text{ mg/dl/min}$ (-0.06 mmol/l/min) y con predicción lineal de que BG cruzará $< 70 \text{ mg/dl}$ dentro de 30 minutos a la RoC actual, y 4) la región de alerta alta (*alert high*) represente $\geq 7.5\%$ y se define como $\text{BG} \leq 300 \text{ mg/dl}$ con $\text{RoC} > +1.5 \text{ mg/dl/min}$ ($+0.08 \text{ mmol/l/min}$) y con predicción de que BG excederá $> 250 \text{ mg/dl}$ (13.9 mmol/l) dentro de 30 minutos a la RoC actual. Todo lo demás constituye la región «neutral», que debería ser $\leq 70\%$ de los pares BG-RoC³².

La pregunta sobre «cómo estructurar la distribución» también implica cómo inducir perfiles glucémicos que la produzcan de manera ética, segura y reproducible. En la propuesta metodológica, la manipulación glucémica se organiza en torno a un *target glucose profile* (TGP) diseñado para atravesar todas las regiones del DGR en una sola sesión y con un periodo de muestreo frecuente (FSP, por sus siglas en inglés) de alrededor de 7 horas, con episodios hiperglucémicos e hipoglucémicos «distintos y transitorios» similares a los observados en vida real³².

La segunda mitad del procedimiento apunta a garantizar la representación de descensos rápidos con hipoglucemia inminente y de hipoglucemia establecida, sin comprometer la seguridad. Tras

alcanzar el pico, se anticipa un descenso rápido hacia la hipoglucemia, atravesando «*alert low*» y «*BG low*»; si no ocurre, se permiten bolos adicionales de insulina. Cuando BG < 70 mg/dl se alcanza o está inminente, se implementan intervenciones de rescate según criterio del equipo y del participante, pero con una recomendación explícita: elegir una estrategia que mantenga BG < 70 mg/dl durante aproximadamente 30 minutos, evitando la hiperglucemia de rebote por ingesta excesiva de carbohidratos. Para el resto del FSP, se sugiere mantener BG estable entre 100 y 180 mg/dl para facilitar el egreso seguro; se recomiendan colaciones (*snacks*)/comidas adicionales de bajo contenido de carbohidratos y considerar la acción residual del primer bolo al tomar decisiones de dosis y tiempos. Si el perfil no evoluciona como se espera y no se alcanzan todas las regiones críticas, se permiten acciones correctivas (siempre seguras), como bolos adicionales, ajuste de la basal en las bombas, ingesta de alimentos o ejercicio leve³².

Aun con un procedimiento estandarizado, la variabilidad biológica y conductual hace improbable alcanzar una distribución «perfecta» sin mecanismos de ajuste, por lo que la estructura recomendada contempla metas ideales y mínimos prácticos. El TGP se diseñó para que, idealmente, alrededor del 11% de los puntos estuvieran en cada región crítica (lo que dejaría ~57% en neutral), pero se propone como criterio mínimo que cada región crítica tenga $\geq 7.5\%$, permitiendo que la región neutral sea como máximo del 70%^{23,24,32,33}.

PREGUNTA 3. ¿QUÉ CRITERIOS DEBEN CONSIDERARSE PARA LA SELECCIÓN Y EL USO DEL DISPOSITIVO DE MONITOREO CONTINUO DE GLUCOSA EN ESTUDIOS CLÍNICOS?

Texto de soporte y análisis

La monitorización continua de glucosa ha pasado de ser una herramienta centrada en el autocuidado de la diabetes a constituir, en investigación

clínica, una plataforma de medición capaz de capturar de manera densa la exposición glucémica (tendencias, variabilidad y episodios transitorios) con un nivel de detalle temporal y de resolución (granularidad) que no es alcanzable con HbA1c ni con la automonitorización capilar aislada. En consecuencia, la selección del sistema debe estar subordinada a la pregunta de investigación y al rol que la MCG desempeñará en el ensayo (intervención, cointervención, herramienta diagnóstica o instrumento de medición de desenlaces). En esta lógica, el consenso internacional sobre MCG y métricas para ensayos clínicos enfatiza la necesidad de planificar explícitamente el uso del dispositivo, definir el objetivo de la medición y alinear los requisitos de datos (ventanas de captura, suficiencia, procesamiento y reporte) con el objetivo del estudio^{36,37}.

Desde la perspectiva de diseño, el primer criterio operativo para elegir y usar un dispositivo de MCG en un estudio clínico consiste en determinar si la señal de MCG será «observacional» (capturada para análisis sin retroalimentación al usuario) o «reactiva» (retroalimentación en tiempo real o intermitente que permite cambios de conducta o de tratamiento). El consenso de métricas para MCG en estudios propone que el diseño especifique: a) si el participante y/o el equipo clínico tendrán acceso a datos en tiempo real; b) si las alertas estarán activas; c) si el dispositivo requiere escaneo (y con qué frecuencia mínima para evitar huecos), y d) si la MCG se implementará como «profesional» (descarga retrospectiva) para reducir la reactividad y estandarizar la exposición. El consenso de ensayos clínicos subraya que las metas del uso de MCG deben declararse explícitamente^{37,38}. Un segundo criterio cardinal es el encuadre regulatorio y de desempeño mínimo exigible para el «uso pretendido» del dispositivo en investigación, especialmente cuando la MCG informará sobre decisiones de dosificación o se integrará con sistemas automatizados. En el contexto de la FDA (Food and Drug Administration), la categoría de monitores continuos de glucosa integrados (iCGM) se asocia con requerimientos de exactitud más estrictos y con la expectativa de que el desempeño esté documentado en

literatura revisada por pares; el reporte sobre iCGM resalta que los sistemas designados iCGM deben cumplir requisitos de desempeño en condiciones ambulatorias que incluyen, entre otros, límites de confianza inferiores del 95% para criterios de exactitud en hipoglucemia, así como restricciones sobre discordancias en RoC en rangos bajos^{36,38}.

En tercer lugar, la selección debe fundamentarse en la validez analítica del dispositivo en el contexto fisiológico y clínico del ensayo, evitando reducir la discusión a una sola métrica como la MARD. Un análisis sistemático sobre la precisión y la confiabilidad de sistemas de MCG en entornos ambulatorios reportó una MARD promedio cercana al 9.4% para los sistemas evaluados, pero también destacó que la exactitud disminuye en hipoglucemia y que ningún dispositivo cumple de forma uniforme criterios estrictos en ese rango, lo que refuerza la necesidad de estrategias de confirmación cuando se sospechan lecturas bajas o discordantes. En paralelo, la carta crítica sobre MARD enfatiza que MARD no es un «estándar regulatorio», que depende de la distribución de glucosas muestreadas y del diseño del protocolo (p. ej., densidad de puntos en los extremos), y que su interpretación aislada puede conducir a conclusiones erróneas si no se acompaña de evaluaciones por rango, por tasa de cambio y por métricas clínicas (p. ej., *error grid*). En consecuencia, un criterio metodológico para seleccionar y usar MCG en ensayos consiste en exigir evidencia de desempeño desagregada por rangos (hipoglucemia, euglucemia, hiperglucemia), por condiciones dinámicas (cambios rápidos), por día de uso del sensor (p. ej., primeras 24-48 h) y por población (edad, comorbilidades)^{24,39}.

La validez analítica debe complementarse con criterios explícitos de «evaluabilidad» del desempeño, porque la literatura muestra una heterogeneidad metodológica que limita la comparabilidad entre dispositivos y estudios. La revisión de alcance sobre recomendaciones para reportar evaluaciones de desempeño de MCG identificó variabilidad sustantiva en diseño (incluyendo elección del método de referencia, distribución de glucosa de las muestras, y reporte de

características del comparador), y subrayó que, con frecuencia, no se reporta adecuadamente la evaluación del método de referencia (p. ej., precisión del analizador, trazabilidad, condiciones preanalíticas), lo cual puede sesgar estimaciones de error atribuibles al sensor. De modo convergente, el informe sobre iCGM enfatiza que la evaluación debe abarcar el rango clínicamente relevante (especialmente < 70 mg/dl) y la tasa de cambio, incluyendo especificación de criterios cuantitativos de concordancia, lo que implica que en ensayos donde el dispositivo es parte del desenlace (p. ej., tiempo bajo rango) o parte del mecanismo (p. ej., AID), la selección debe privilegiar dispositivos con documentación pública, protocolos de validación robustos y métricas de desempeño más allá de promedios globales^{23,36}.

La pertinencia de estos criterios se evidencia al observar que los efectos clínicos atribuidos a «usar MCG» dependen de la tecnología concreta seleccionada y de su implementación. En adultos con DT1, un metaanálisis de 27 estudios (2,975 participantes) encontró que el uso de MCG se asoció con reducción de HbA1c (DM: -0.38%, IC 95%: -0.49 a -0.27) y aumento del TIR (DM: +7.9%, IC 95%: 5.8-10.0), con reducción de la exposición a hipoglucemia³³.

Cuando el objetivo del ensayo requiere comparar tecnologías de MCG (en lugar de compararlas con tecnologías no MCG), la evidencia sintetizada muestra diferencias sistemáticas atribuibles a las características funcionales de los dispositivos. Un metaanálisis que comparó MCG en tiempo real vs. MCG intermitente por escaneo en DT1 reportó un incremento de TIR de +7.0 puntos porcentuales (IC 95%: 5.8-8.3; I^2 : 0), y una reducción de TBR < 70 mg/dl de -1.7 (IC 95%: -2.5 a -0.8), sin diferencia significativa en HbA1c. De manera complementaria, un metaanálisis de rtCGM vs. *flash glucose monitoring* (FGM) encontró mejoras pequeñas pero consistentes: reducción de HbA1c (diferencia de medias estandarizada [DME]: -0.20, IC 95% -0.33 a -0.08), incremento de TIR (DME: +0.19, IC 95% 0.06-0.32), reducción de TAR (DME: -0.22, IC 95%: -0.35 a -0.09) y reducción de TBR (DME: -0.11, IC 95% -0.23 a 0.00)^{34,40}.

En la DT2 no tratada con insulina, la selección del dispositivo y su modo de uso adquieren una dimensión adicional: la MCG puede operar como intervención educativa/comportamental y, por ello, alterar el comparador habitual. Una RS con metaanálisis de 15 ensayos ($n = 2,506$) concluyó que la MCG se asoció con una reducción de HbA1c (diferencia de medias ponderada: -0.37% , IC 95%: -0.50 a -0.24), un aumento de TIR ($+8.84$ puntos porcentuales, IC 95%: 5.73 - 11.94) y una reducción de TAR (-8.14 , IC 95%: -11.25 a -5.04). En paralelo, una RS con metaanálisis de ECA en DT2 encontró que la MCG mejoró la HbA1c (DM 0.32% a favor de MCG; con análisis por subgrupos, donde el efecto fue más consistente para rtCGM que para FGM), aumentó la satisfacción y reportó un incremento en eventos adversos (RR: 1.30 , IC 95%: 1.08 - 1.57), lo que obliga a incorporar consideraciones de seguridad y tolerabilidad en el plan de seguimiento⁴¹⁻⁴³.

Un metaanálisis sobre MCG profesional (14 ensayos, 2,451 participantes) reportó mejoras en HbA1c (DM: -0.28% , IC 95%: -0.36 a -0.20) y en TIR ($+5.59$ puntos porcentuales, IC 95%: 3.32 - 7.87), con una reducción de TAR (-5.39 , IC 95%: -7.49 a -3.30) y una reducción menor de TBR < 70 mg/dl (-0.45 , IC 95%: -0.83 a -0.06). Estos resultados son relevantes para la selección del dispositivo en ensayos porque sugieren que incluso modalidades retrospectivas pueden asociarse a mejoras clínicas, probablemente mediadas por ajustes terapéuticos informados por patrones, pero la interpretación depende críticamente de si el protocolo permite usar los datos para modificar el tratamiento durante la fase de evaluación del desenlace^{37,44}.

En ensayos de tecnologías AID, la selección del dispositivo de MCG adquiere carácter estructural, ya que la MCG no solo mide, sino que también alimenta el algoritmo que determina microbolos, la suspensión o los ajustes basales. Una RS con metaanálisis en red de sistemas híbridos de asa cerrada (HCL) mostró que, globalmente, la HCL incrementa el TIR en aproximadamente $+16.56$ puntos porcentuales (IC 95%: 10.55 - 22.58) y reduce la HbA1c (DM: -0.41% , IC 95%: -0.62 a -0.21) frente a terapia convencional o sistemas

menos automatizados; además, al comparar sistemas específicos, algunos (p. ej., MiniMed 780G) presentaron el mayor incremento de TIR ($+18.0$, IC 95%: 12.1 - 23.9) y una mejora significativa en TBR (-1.84 , IC 95%: -2.74 a -0.94) vs. *sensor-augmented pump* (SAP)⁴⁵.

Un criterio transversal, independientemente del área terapéutica, es la estandarización de las métricas y la suficiencia de los datos. El consenso internacional sobre uso de MCG recomienda que, para interpretar perfiles glucémicos de forma confiable, se utilicen al menos 14 días de datos con $\geq 70\%$ de registros esperados; además, promueve la adopción de métricas centrales como TIR (70-180 mg/dl en la mayoría de los adultos con diabetes), TBR (con cortes < 70 y < 54 mg/dl), TAR, glucosa media, variabilidad (p. ej., coeficiente de variación) y, cuando sea pertinente, indicadores derivados (p. ej., indicador de manejo de glucosa), resaltando que estas métricas capturan dimensiones distintas y no son intercambiables con HbA1c^{37,38,46}.

La dimensión analítica se vuelve más compleja porque, aun con el mismo archivo de datos, diferentes paquetes estadísticos, algoritmos de filtrado y reglas de manejo de datos faltantes pueden producir resultados no idénticos; esto impacta directamente en la comparabilidad entre ensayos y en la replicabilidad. Una RS sobre paquetes y algoritmos para el análisis de datos de MCG identificó una diversidad de herramientas y reportes, y señaló que los informes de MCG varían entre fabricantes y que los investigadores utilizan distintos métodos de procesamiento, lo que puede generar discrepancias en las métricas calculadas y en la presentación de resultados; en este contexto, la revisión resalta la necesidad de estandarización de los reportes y de transparencia en los algoritmos^{28,38}.

Los criterios de selección y uso no pueden limitarse al desempeño y a la analítica; la factibilidad humana determina la integridad de los datos, su retención y su generalización. La revisión sobre barreras y oportunidades para MCG en adultos con DT2 documentó facilitadores (mejor entendimiento de patrones, apoyo a la toma de

decisiones, sensación de control) y barreras recurrentes (costos, alfabetización digital, carga de aprendizaje y preocupaciones sobre la precisión) y, además, ilustró cómo las preocupaciones sobre la precisión (incluidos los valores de MARD reportados para dispositivos específicos) pueden erosionar la confianza y reducir el uso sostenido⁴⁷⁻⁴⁹.

Finalmente, la selección y el uso del dispositivo en ensayos deben considerar contextos clínicos en los que la exactitud puede degradarse y en los que la seguridad exige redundancia de medición. En el ámbito perioperatorio de cirugía no cardíaca, una RS reportó variabilidad amplia de exactitud (p. ej., medianas de MARD entre 9.0 y 22.0% según estudio y tecnología), con heterogeneidad en métodos y poblaciones, lo que sugiere que extrapolar desempeño ambulatorio al entorno quirúrgico es metodológicamente riesgoso; en tales contextos, un criterio razonable es seleccionar dispositivos con mejor documentación en ese escenario o, si no existe, incorporar confirmación capilar/arterial como estándar de seguridad y como referencia para adjudicación de eventos. De forma coherente, los estándares clínicos de tecnología en diabetes enfatizan que, aun cuando la MCG se use de manera sistemática, la medición capilar sigue siendo necesaria como respaldo en situaciones de discrepancia, cambios rápidos o sospecha de error del sensor, y esta lógica es transferible al diseño de ensayos para prevenir desenlaces de seguridad mal clasificados y reducir el sesgo de medición^{26,36-38}.

PREGUNTA 4. ¿CUÁLES SON LAS MÉTRICAS RECOMENDADAS PARA EVALUAR LA PRECISIÓN Y LA EXACTITUD DE LOS SISTEMAS DE MONITOREO CONTINUO DE GLUCOSA EN ESTUDIOS CLÍNICOS?

Texto de soporte y análisis

La evaluación del desempeño de los sistemas MCG en estudios clínicos debe entenderse

como un ejercicio de validez metrológica y de seguridad clínica: no basta con demostrar que un sensor «correlaciona» con una glucosa de referencia o que mejora desenlaces glucémicos promedio, sino que es imprescindible caracterizar, con métricas estandarizadas y reproducibles, tanto la exactitud (cercanía a la referencia) como la precisión (repetibilidad/reproducibilidad) en todo el rango glucémico y bajo condiciones dinámicas (cambios rápidos), porque son precisamente esos escenarios los que determinan el riesgo de decisiones terapéuticas erróneas, fallas de alertas y desempeño subóptimo en interoperabilidad con algoritmos y dispositivos conectados. En una revisión del alcance de las evaluaciones clínicas de desempeño de MCG (2002-2022), se identificó una heterogeneidad sustancial en los métodos y parámetros reportados, y se subrayó que la comparabilidad entre estudios solo es sólida cuando los sistemas se evalúan en el mismo estudio o bajo protocolos estandarizados^{23,36}.

Desde una perspectiva de las ciencias de la medición aplicada a la tecnología médica, la exactitud y la precisión son constructos distintos y complementarios: la exactitud se relaciona con el sesgo (*bias*) y el error sistemático respecto a una referencia trazable, mientras que la precisión se relaciona con el error aleatorio (variación entre mediciones) y su estabilidad bajo condiciones repetidas o paralelas. En el ámbito regulatorio de las pruebas diagnósticas, se ha insistido en que los dos parámetros más relevantes para el desempeño son precisamente el sesgo y la precisión, con metodologías típicamente alineadas con protocolos de estandarización (p. ej., familias de procedimientos tipo protocolo de evaluación [EP] para sesgo y precisión), porque un dispositivo puede exhibir un sesgo pequeño pero una precisión deficiente (o viceversa), y ambos perfiles implican riesgos clínicos diferentes^{23,39}.

En la práctica clínica y regulatoria, la métrica numérica más difundida para el error puntual es la diferencia relativa absoluta (ARD, *absolute relative difference*) y su agregación como MARD,

que resume la magnitud promedio del error relativo absoluto entre lecturas de MCG y valores de referencia emparejados. En consecuencia, se recomienda reportar no solo una MARD global, sino también su distribución (p. ej., por sensor) y medidas de dispersión, así como intervalos de confianza, ya que los criterios regulatorios contemporáneos para MCG integrados se basan explícitamente en límites inferiores unilaterales del 95% para porcentajes de lecturas dentro de tolerancias predefinidas^{23,36}.

La recomendación metodológica actual también enfatiza que el desempeño debe describirse de manera estratificada por rangos glucémicos clínicamente relevantes, dado que el comportamiento del error cambia debido a no linealidades fisiológicas, limitaciones electroquímicas y algoritmos de filtrado/retardo, y porque el significado clínico de un error relativo es diferente a bajas concentraciones. En una RS centrada en la precisión y la confiabilidad en entornos ambulatorios, se documentó que los estudios emplean umbrales diversos para hipoglucemia e hiperglucemia y que el rendimiento en hipoglucemia puede ser marcadamente peor y más variable^{24,36}.

Complementariamente, se recomiendan las tasas de concordancia, que cuantifican el porcentaje de lecturas de MCG dentro de límites absolutos o relativos respecto a la referencia (p. ej., ± 15 mg/dl o $\pm 15\%$, $\pm 20/20$, $\pm 40/40$). La revisión de alcance identificó que las tasas de acuerdo se reportan en una proporción importante de estudios y que cuando se calcula una tasa única en todo el rango es frecuente cambiar de diferencia relativa a absoluta por debajo de un umbral (para evitar la inflación del error relativo en glucosas bajas), lo cual debe describirse explícitamente para permitir comparaciones válidas^{23,36,38}.

No obstante, ni MARD ni las tasas de concordancia capturan por sí solos el impacto clínico diferencial de un error (no es equivalente errar 25 mg/dl cerca de 55 mg/dl que errar cerca de 250 mg/dl), razón por la cual se recomienda incorporar métricas de exactitud clínica basadas en análisis de rejillas de error (*error grids*). En la

revisión de alcance, el análisis por rejilla fue el enfoque predominante para la exactitud clínica y se reconoció como el único método «establecido» para traducir las desviaciones MCG-referencia en zonas de impacto clínico, aunque coexisten varias rejillas (Clarke, Consensus/Parkes, Surveillance y CG-EGA). La coexistencia de múltiples rejillas y variaciones en su implementación limita la comparabilidad, lo que ha impulsado recomendaciones recientes hacia el uso de rejillas ampliamente aceptadas y sensibles al riesgo contemporáneo; en particular, la *consensus error grid* ha ganado terreno en años recientes y se promueve como parte de una evaluación multidimensional junto con MARD y tasas de concordancia^{23,24,38}.

Una dimensión frecuentemente subestimada, pero crítica para MCG actualmente (especialmente con alertas predictivas e interoperabilidad), es la exactitud de tendencia (*trend accuracy*), que evalúa la capacidad del sistema para reflejar la dirección y la magnitud de los cambios (subidas/bajadas), no solo los puntos estáticos. La revisión de alcance documentó que cuando se reporta *trend accuracy*, un componente común ha sido el *rate error grid* (parte de CG-EGA) y, además, se han utilizado enfoques como el cálculo de desviaciones de tasa (análogas a desviaciones puntuales) y tablas de concurrencia; sin embargo, el reporte fue minoritario y metodológicamente heterogéneo, por lo que se recomienda estandarizarlo^{36,38}.

Para que la *accuracy* de tendencia y, en general, la exactitud observada, sean interpretables, se recomienda reportar métricas que describan el contexto dinámico de la evaluación, particularmente la distribución de tasas de cambio de la referencia; en la revisión de alcance se propuso reportar un parámetro resumen, como el MARoC de la glucosa de referencia, acompañado de una descripción adecuada de la distribución de RoC, para capturar cuán exigente fue el escenario dinámico^{36,38}.

En paralelo a la exactitud frente a la referencia, la precisión del MCG debe cuantificarse explícitamente porque aporta información que no

se recupera de MARD: un sistema puede exhibir buena MARD promedio por compensación de errores o por efectos del diseño del estudio, pero mantener alta variabilidad aleatoria entre sensores o a lo largo del tiempo, lo que afecta la estabilidad de las alertas, la variabilidad estimada y la seguridad en la automatización. Un planteamiento clave es que la caracterización del desempeño mejora cuando se usan métricas conjuntas de exactitud y precisión; específicamente, se ha propuesto que el uso combinado de MARD (exactitud respecto a la referencia) y PARD (*precision absolute relative deviation*, calculada como la diferencia relativa absoluta entre dos sensores idénticos usados en paralelo) permite una caracterización más completa del desempeño del sensor^{38,51}.

Una extensión natural de la evaluación de precisión es la cuantificación de la variabilidad sensor-a-sensor, que se recomienda reportar mediante la distribución de métricas por sensor (p. ej., la distribución de MARD individual por sensor o la dispersión de PARD en pares), porque la seguridad clínica no está determinada solo por el promedio poblacional, sino por la cola de peor desempeño y su frecuencia. La revisión de alcance resaltó que la variabilidad entre sensores es un aspecto «muy importante» de la exactitud puntual y puede caracterizarse mediante la distribución de MARD individuales, lo que permite inferir prácticas sobre la consistencia de fabricación, la estabilidad del algoritmo a lo largo de su vida útil y el riesgo de «sensores atípicos» con desempeño deficiente^{23,51}.

El retardo (*lag*) entre la glucosa intersticial estimada y la glucosa sanguínea de referencia es otra dimensión que debe medirse y reportarse, porque sesga de forma diferencial la exactitud en cambios rápidos y afecta tanto la exactitud de tendencia como el desempeño de las alertas⁵¹.

Para estudios con objetivos clínicos centrados en la seguridad (hipoglucemia/hiperglucemia) y para tecnologías con alertas, se recomiendan métricas de desempeño de tipo diagnóstico, especialmente la sensibilidad y la especificidad

(y, cuando sea posible, los valores predictivos y las tasas de falsas alarmas) para la detección de eventos por debajo o por encima de umbrales^{24,36}.

La evaluación recomendada también incluye métricas de confiabilidad técnica y de disponibilidad de datos, ya que un MCG clínicamente «exacto» pero con brechas frecuentes o transmisión inestable puede provocar fallas de interoperabilidad, sesgos en métricas derivadas (TIR/TBR/TAR, variabilidad) y eventos adversos por pérdida de alertas. En el marco de MCG integrados, se exige demostrar la ausencia de brechas clínicamente significativas que impidan a los dispositivos conectados lograr su uso previsto, y se solicita describir la frecuencia y la duración de las brechas, además de la vida útil observada del sensor^{23,36,38}.

Las recomendaciones contemporáneas convergen, por tanto, hacia un enfoque multidimensional donde la «exactitud» se describe con al menos tres familias de métricas: a) error numérico puntual (ARD/MARD, preferentemente con distribución e intervalos de confianza y estratificación por rangos); b) exactitud clínica (rejillas de error, reportando proporciones por zonas), y c) tasas de acuerdo (porcentajes dentro de tolerancias 15/15, 20/20, 40/40, con umbrales de cambio absoluto/relativo explícitos y con límites inferiores de confianza cuando aplique)^{23,36,38}.

La importancia de estandarizar y reportar estas métricas de desempeño del dispositivo se vuelve más evidente cuando se observa que los desenlaces «clínicos» basados en MCG (TIR/TBR/TAR, variabilidad) muestran beneficios consistentes en metaanálisis, pero su magnitud y comparabilidad dependen de la integridad y la exactitud de los datos del sensor. En una RS con metaanálisis, el uso de MCG frente al monitoreo estándar se asoció con una reducción promedio de HbA1c (DM: -0.38%), un incremento absoluto de TIR (~7.9%) y una reducción de la exposición a hipoglucemia; y, en comparaciones directas rtCGM vs. isCGM se observaron diferencias a favor de rtCGM en HbA1c (-0.23%) y TIR (+5.36%)^{26,40}.

PREGUNTA 5. ¿CUÁL ES EL PAPEL DE LA CALIBRACIÓN MANUAL CAPILAR EN LA PRECISIÓN DEL MONITOREO CONTINUO DE GLUCOSA?

Texto de soporte y análisis

En el marco regulatorio de dispositivos diagnósticos, la «precisión» (repetibilidad/reproducibilidad; error aleatorio) y el «sesgo» (error sistemático) son parámetros nucleares para caracterizar el desempeño analítico y comparativo, y su interpretación depende críticamente de cómo se obtienen y utilizan los datos de referencia y de ajuste del sistema³⁹. Todos los sistemas de MCG requieren algún tipo de calibración para mantener la exactitud clínica a lo largo del tiempo de uso del sensor, la cual puede realizarse por el propio usuario (calibración manual con glucosa capilar) o de manera predefinida en fábrica (calibración de manufactura)^{36,38} (Tabla 5).

En términos de mecánica, la calibración manual capilar funciona como una operación de «anclaje» del algoritmo del sensor: el sistema ajusta la transformación de la señal del sensor (glucosa intersticial estimada) a una concentración reportada en mg/dl mediante puntos de calibración provenientes de una medición capilar. En teoría, esto puede corregir la deriva del sensor y reducir el sesgo (mejorando métricas de exactitud como MARD o porcentajes dentro de criterios 15/15, 20/20), pero en la práctica desplaza parte del desempeño del MCG hacia la calidad del procedimiento de punción capilar, del medidor (BGMS) y de la introducción correcta del valor^{36,38}.

La relevancia de la calibración manual capilar para la precisión del MCG se entiende mejor al reconocer que el «calibrador» (glucosa capilar) no es una verdad absoluta, sino un dato con un error propio, susceptible de variaciones preanalíticas y analíticas. En una RS sobre la evaluación clínica del desempeño de MCG y recomendaciones de reporte, se documenta que la glucosa capilar (medida típicamente por punción digital) fue utilizada históricamente para decisiones

terapéuticas y también se utiliza para la calibración de sistemas MCG; incluso en sistemas no adjuntos (*no-adjunctive*), las mediciones capilares se recomiendan cuando los valores del MCG son implausibles o temporalmente no están disponibles²³.

La calibración manual capilar introduce variabilidad adicional porque la medición capilar en sí misma depende de las condiciones de toma de muestra y del desempeño del medidor y de las tiras reactivas; esta variabilidad se convierte en variabilidad del MCG cuando se usa como insumo para calibrar. La misma revisión de alcance discute que la repetición de punciones digitales durante muestreos frecuentes es una desventaja operativa (fatiga/adhesión) y resalta que la inexactitud del comparador por BGMS (parcialmente por errores preanalíticos) es un problema porque, en la lectura del estudio, se «carga» al MCG. Las GPC tituladas Estándares de Cuidado en Diabetes 2025 de la Asociación Americana para la Diabetes (ADA, American Diabetes Association) señalan que los usuarios de MCG deben mantener acceso a monitoreo capilar (BGM), entre otros motivos, para situaciones de sospecha de inexactitud del sensor, durante periodos de calentamiento, pérdida de señal y para calibración si el dispositivo lo requiere, reforzando que la calibración manual no sustituye sino que incrementa la dependencia operativa del dato capilar^{16,23}.

Un componente fisiológico que condiciona directamente el impacto de la calibración manual en la precisión observada es la discordancia temporal/compartimental entre la glucosa intersticial (lo que capta el sensor) y la glucosa sanguínea capilar (lo que se usa para calibrar). En los estándares ADA 2025 se describe que pueden existir diferencias entre valores del MCG y mediciones capilares durante periodos de cambios rápidos de glucosa (p. ej., > 2 mg/dl/min) o durante las primeras 12 horas tras la inserción del sensor; esto es especialmente relevante porque una calibración realizada durante un cambio rápido puede «forzar» al algoritmo a alinear dos compartimentos que, por dinámica fisiológica, no están sincronizados, con el riesgo de introducir error

Tabla 5. Calibración en sistemas de MCG: impacto en desempeño, riesgos, mitigación y consideraciones para ensayos clínicos

Tipo de calibración	Ventajas metrológicas	Riesgos típicos	Recomendaciones de práctica segura (cuándo calibrar/cuándo confirmar con BGM)	Implicación para ensayos clínicos (estandarización, entrenamiento, documentación)
Calibración manual (capilar)	– Ancla el algoritmo de calibración del sistema a mediciones capilares (BGM), permitiendo ajustar el desempeño del sensor conforme al procedimiento definido por el fabricante³⁸	– Error analítico del glucómetro capilar (BGM) utilizado como calibrador; el error del BGM se transfiere al sensor calibrado^{16,38} – Errores de usuario: introducción incorrecta del valor de calibración y/o calibración en momentos fisiológicamente inadecuados; estos errores pueden incrementar la MARD observado en práctica³⁸ – Dependencia del protocolo y del algoritmo de calibración específico del dispositivo, lo que dificulta la comparabilidad entre estudios si no se estandariza y reporta de forma explícita^{5,46}	– Asegurar disponibilidad de BGM para calibración (cuando aplique) y para confirmación de lecturas cuando exista discrepancia clínica¹⁶ – Confirmar con BGM ante tasas de cambio rápidas (p. ej., > 2 mg/dl/min) y/o cuando los síntomas no concuerden con el valor del MCG¹² – En evaluaciones de desempeño, preferir BGM capilar como respaldo operativo frente a valores de laboratorio para el uso clínico, según consenso internacional³⁸	– Definir y reportar el protocolo de calibración (frecuencia, procedimiento), incluyendo fuente y concentración de los calibradores cuando corresponda⁵ – Documentar explícitamente si el sistema requiere calibración manual (o si es opcional) y describir el algoritmo/método de calibración empleado para evitar sesgos de comparabilidad entre brazos, sitios y dispositivos^{38,46} – Estandarizar entrenamiento y verificación de competencia del usuario/investigador; registrar desviaciones del protocolo y eventos de seguridad relacionados^{5,38}
Calibración de fábrica (<i>factorycalibrated</i> , típicamente iCGM)	– Elimina la necesidad de calibración por parte del usuario, reduciendo carga operativa y variabilidad inducida por errores humanos (digitación/calibración inadecuada)^{36,38} – Alinea el desempeño con la expectativa regulatoria de exactitud para sistemas interoperables (iCGM) y soporta uso no adyuvante conforme a etiquetado³⁶	– Persisten factores fisiológicos y exógenos que pueden afectar la exactitud (p. ej., hematocrito, deshidratación, compresión del sitio de inserción, interferentes), aun sin calibración manual³⁸ – Riesgo de confianza excesiva si no se confirma con BGM cuando hay discordancia clínica o cambios rápidos de glucosa^{16,38}	– Mantener acceso a BGM como método de respaldo y confirmación¹⁶ – Confirmar con BGM en tasas de cambio rápidas (> 2 mg/dl/min), discrepancia con síntomas o sospecha de fallo/artefacto del sensor^{16,38} – Educar al usuario sobre interferencias/condiciones que degradan exactitud y sobre acciones de mitigación (p. ej., repetir lectura, verificar sitio)³⁸	– Reportar explícitamente que el dispositivo es de calibración de fábrica; documentar fuente del dispositivo/sensores, lotes y vida útil planificada para trazabilidad⁵ – Especificar procedimientos de confirmación con BGM y criterios de exclusión de datos por fallos técnicos para preservar validez interna^{5,16} – Cuando el objetivo sea interoperabilidad y/o desempeño tipo iCGM, alinear diseño y reporte con métricas/expectativas descritas en documentos de consenso y literatura sobre estándares FDA^{36,38}

BGM/SMBG: monitoreo de glucosa en sangre por glucómetro (capilar); iCGM: CGM integrable, *integrated continuous glucose monitoring*; MARD: *mean absolute relative difference*, diferencia relativa absoluta media; MCG/CGM: monitoreo continuo de glucosa. Adaptada de Freckmann et al.⁵, ElSayed et al.¹⁶, Klonoff et al.³⁶, Scharf, et al.³⁸, y Danne et al.⁴⁶.

sistemático y variabilidad posterior (peor precisión a lo largo del tiempo)^{16,23}.

En una RS sobre precisión y confiabilidad de sistemas rtCGM en entornos ambulatorios, se documentan protocolos de «mínima calibración» (p. ej., 2 veces/día) y «calibración adicional» (p. ej., 3-4 veces/día) en estudios con ciertos sensores, y se reporta, como ejemplo, que para un sistema (Guardian Sensor 3 con MiniMed 640G) el porcentaje dentro de 20 mg/dl o 20% en un umbral de 70 mg/dl fue 92.5% con calibración mínima y 92.8% con calibración adicional. Este tipo de diferencia pequeña sugiere que, bajo condiciones controladas y con procedimientos estandarizados, aumentar las calibraciones no necesariamente transforma de manera sustantiva el desempeño puntual^{16,23,24}.

PREGUNTA 6. EN DISPOSITIVOS QUE REQUIEREN CALIBRACIÓN MANUAL, ¿CUÁL ES LA MEJOR ESTRATEGIA DE CALIBRACIÓN?

Texto de soporte y análisis

El consenso internacional sobre el uso de MCG enfatiza explícitamente que, a diferencia del iCGM retrospectivo, la exactitud de muchos rtCGM depende de pruebas capilares (SMBG) para la calibración y que, en consecuencia, se requiere un glucómetro preciso y una ejecución correcta del procedimiento; además, señala que la calibración exitosa depende de condiciones específicas, entre ellas realizarla cuando la glucosa no cambia rápidamente y capacitar al usuario en técnicas apropiadas⁴⁶.

Desde la ciencia de la medición, la calibración manual debe analizarse bajo un modelo de «error compuesto»: el MCG aporta un componente de error propio (sensor + algoritmo) y la medición capilar añade un error adicional (preanalítico, analítico y humano) que, al incorporarse como punto de calibración, puede propagarse a múltiples lecturas posteriores. En este marco, calibrar

con un valor capilar ruidoso o sesgado puede reducir la exactitud aparente a corto plazo o aumentar la variabilidad longitudinal del error, porque el ajuste algorítmico incorpora un dato que no es un «valor verdadero» sino una medición con incertidumbre; por ello, la calidad de la calibración depende tanto del sensor como del calibrador. En un análisis centrado en requisitos de desempeño para la evaluación comparativa de tecnologías (incluyendo marcos del Instituto de Estándares Clínicos y de Laboratorio [CLSI, Clinical and Laboratory Standards Institute], como EP5 para precisión y EP9 para comparación de métodos), se resalta que la precisión y el sesgo son parámetros cardinales y que cualquier procedimiento que modifique el sesgo (como la calibración) debe ejecutarse de manera controlada para no introducir nueva variabilidad⁴⁶.

En una revisión sobre evaluación clínica del desempeño de MCG y recomendaciones de reporte, se documenta que los sistemas calibrados en fábrica se han vuelto cada vez más prevalentes, pero que la calibración continúa siendo un tema mayor en la exactitud del MCG; en la mayoría de los estudios revisados, la calibración se realizó siguiendo instrucciones del fabricante y con frecuencia no se describió con suficiente detalle, mientras que solo unos pocos trabajos exploraron explícitamente el efecto de calibraciones adicionales y/o el momento de calibrar sobre la exactitud²³.

El consenso internacional de MCG afirma explícitamente que la calibración se realiza mejor cuando la glucosa no varía rápidamente. De manera concordante, las GPC ADA 2025 subrayan que las personas que utilizan MCG deben mantener acceso a BGM por múltiples razones que incluyen el periodo de inicialización (*warm up*), interrupciones de transmisión, la necesidad de calibración (si aplica) y escenarios clínicos donde la glucosa cambia rápidamente (> 2 mg/dl/min); aunque el texto se centra en la indicación de confirmación/seguridad, su implicación metrológica para calibración manual es directa: calibrar en momentos de alta tasa de cambio incrementa el riesgo de «anclar» el sensor a una

referencia que no representa el mismo compartimento/tiempo fisiológico^{16,46}.

Un segundo pilar de la estrategia óptima es controlar la calidad del calibrador capilar, ya que la calibración manual, por definición, transfiere al MCG la incertidumbre del glucómetro y del procedimiento de medición. El consenso internacional establece que, si el rtCGM depende del SMBG para su calibración, es importante contar con un medidor de glucosa preciso y educar al usuario en técnicas apropiadas, reconociendo que la calibración exitosa requiere condiciones específicas. Las GPC ADA 2025 refuerzan que existen diferencias de exactitud entre glucómetros y que el acceso a BGM sigue siendo necesario en múltiples escenarios operativos; desde el punto de vista de la estrategia de calibración en investigación clínica, esto sugiere buenas prácticas como estandarizar el modelo de glucómetro y las tiras (idealmente provistos por el estudio), reducir la heterogeneidad entre medidores (que puede traducirse en heterogeneidad de calibración) y documentar el procedimiento y la adherencia^{16,46}.

La selección de la estrategia también debe considerar la dirección regulatoria y tecnológica: múltiples documentos recientes enfatizan que la calibración de fábrica reduce la carga del usuario y elimina errores introducidos por la calibración manual, lo cual, por extensión, reduce las fuentes de variabilidad en estudios clínicos. Un reporte enfocado en iCGM remarca que la calibración es necesaria para mantener la precisión y exactitud del MCG y que, cuando está disponible, la calibración de fábrica es preferible porque reduce errores del paciente y la carga operativa; de forma convergente, un consenso latinoamericano sobre la estandarización de métricas de evaluación destaca que los sistemas calibrados en fábrica reducen la carga y potencialmente mejoran la consistencia al minimizar errores asociados al usuario. En dispositivos donde la calibración manual sigue siendo obligatoria, estos argumentos no niegan su uso, pero sí elevan el estándar de «mejor estrategia»: debe incluir capacitación estructurada, provisión/estandarización de

glucómetro, reglas de cuándo calibrar y cuándo confirmar (en especial ante hipoglucemia sospechada), y documentación detallada del proceso en el protocolo del estudio, dado que una calibración inadecuada puede comprometer tanto la seguridad clínica como el cumplimiento de criterios estrictos de exactitud exigidos para interoperabilidad y usos no adyuvantes^{23,36}.

PREGUNTA 7. ¿CUÁL ES LA RECOMENDACIÓN DE ACUERDO CON LOS ESTÁNDARES INTERNACIONALES, LOS CONSENSOS Y LOS ESTÁNDARES ICGM DE LA FDA RESPECTO A LA GENERACIÓN DE REPORTES Y MANEJO DE LA INFORMACIÓN?

Texto de soporte y análisis

La recomendación contemporánea sobre generación de reportes y manejo de la información en MCG^{23,28} parte de un principio metodológico: el MCG es, simultáneamente, un sistema de medición y un sistema de información (captura-transmisión-almacenamiento-procesamiento-visualización), por lo que la validez clínica y regulatoria depende tanto de la exactitud/precisión del sensor como de la trazabilidad, integridad y reproducibilidad de los datos y reportes derivados²³. En términos de estándares internacionales para reportes clínicos, el consenso internacional establece que estandarizar el reporte y el análisis de la glucosa en un formato «tipo electrocardiograma» es vital para optimizar las decisiones clínicas y la comparabilidad entre consultas, centros y contextos. En ese marco, se recomienda la visualización estandarizada mediante el perfil ambulatorio de glucosa (AGP, *ambulatory glucose profile*) como formato de referencia para resumir, de forma interpretable y consistente, la distribución temporal de la glucosa, los patrones recurrentes y las métricas agregadas, reconociendo que existen múltiples plataformas comerciales, pero deben converger en una salida estandarizada. En la lógica de

manejo de la información, esto implica que la «salida» clínica no debe depender de reportes idiosincráticos del fabricante, sino de un conjunto mínimo de visualizaciones y estadísticas comparables y transportables entre sistemas⁴⁶.

La estandarización del reporte no es solo un tema de formato; también exige suficiencia y estabilidad de los datos para que el reporte sea clínicamente interpretable y comparable. El consenso internacional establece que la interpretación retrospectiva de MCG depende de cierta estabilidad de los patrones día a día y propone un umbral operativo: un mínimo de 14 días consecutivos de datos con aproximadamente el 70% de lecturas esperadas en esos 14 días, como base para generar un reporte que habilite el análisis y la toma de decisiones óptimas⁴⁶.

Desde la perspectiva del flujo asistencial y de la gobernanza de datos, el consenso internacional también subraya que la integración de métricas estandarizadas en expedientes clínicos electrónicos es importante para optimizar el flujo de trabajo y facilitar la comunicación remota con los pacientes, lo que desplaza el problema desde «ver una gráfica» hacia «administrar información longitudinal» con trazabilidad y continuidad. En paralelo, las GPC ADA 2025 refuerzan que el uso clínico seguro requiere procedimientos de verificación y respaldo (p. ej., acceso a BGM para discrepancias, periodos de *warm-up* o pérdida de señal), lo cual tiene implicaciones directas de reporte y manejo: los sistemas y reportes deben identificar y documentar periodos sin señal, condiciones de potencial discordancia y eventos operativos que afectan interpretabilidad (p. ej., cambios rápidos, fallas de sensor), en vez de presentar métricas agregadas «sin contexto»^{16, 46}.

En el contexto de la evaluación clínica del desempeño, las recomendaciones del grupo de trabajo asociado a la IFCC (International Federation of Clinical Chemistry and Laboratory Medicine) enfatizan que el reporte debe contener suficiente detalle para permitir una interpretación informada de los resultados y, por ello, proponen un marco explícito de reporte que abarque

el diseño general del estudio, el uso del sistema MCG, la aproximación de medición comparadora (de referencia), los procedimientos de prueba y el análisis estadístico del desempeño²³. Un elemento clave de estas recomendaciones es que la transparencia del reporte debe impedir que la «calidad» del estudio dependa de supuestos no verificables. La revisión de alcance documenta que información crucial sobre el diseño frecuentemente no se reporta, lo cual pone en riesgo la integridad y validez de los hallazgos; por ello, la recomendación práctica es que los reportes incluyan, como mínimo, características de la población, criterios de inclusión/exclusión (incluyendo fármacos interferentes), número de centros, justificación de tamaño de muestra, y detalles del sistema MCG (número de sensores, sitios de inserción, condiciones de uso y calibración si aplica), además de detallar el enfoque del comparador (tipo de muestra, método de laboratorio o dispositivo, trazabilidad y condiciones preanalíticas). En la gestión de la información, esto se traduce en un requisito de metadatos completos y en la necesidad de conservar evidencia verificable del *pipeline* (desde la captura de referencia hasta el emparejamiento temporal y el filtrado), porque la comparabilidad interestudios depende tanto del dato como de sus reglas de construcción²³.

La relevancia de estas recomendaciones se observa indirectamente en la evidencia de metaanálisis: los desenlaces que hoy se consideran clínicamente significativos (HbA1c, TIR/TBR/TAR) dependen de datos de MCG bien gestionados y de reportes comparables. En DT1, una RS con metaanálisis y metaanálisis en red reportó que MCG redujo la HbA1c con una diferencia media de aproximadamente -0.38% (IC 95%: -0.49 a -0.27) e incrementó TIR en un 7.9% (IC 95%: 5.8-10.0), además de ventajas del rtCGM frente al isCGM para TIR (+5.63%) y reducción de la hipoglucemia (-1.28%). En DT2, una revisión con metaanálisis tradicional y en red encontró, frente a SMBG, una reducción de HbA1c (DM: -0.31; IC 95%: -0.45 a -0.18), una reducción de TAR (DM: -9.06%), una reducción de TBR (DM: -0.30%) e incremento de TIR (DM: +8.49%), con

rtCGM como la alternativa con mayor mejora en el metaanálisis en red^{33,51}.

PREGUNTA 8. ¿CUÁL ES LA RECOMENDACIÓN, DE ACUERDO CON LOS ESTÁNDARES INTERNACIONALES, LOS CONSENSOS Y LOS ESTÁNDARES ICGM DE LA FDA, RESPECTO A LA SEGURIDAD DEL USUARIO Y LOS SERVICIOS DE SOPORTE DEL DISPOSITIVO?

Texto de soporte y análisis

En el marco regulatorio contemporáneo, el «servicio al usuario» y la «seguridad del dispositivo» en los sistemas de MCG deben entenderse como componentes inseparables del desempeño clínico: la seguridad depende tanto de la exactitud/fiabilidad metrológica como de la capacidad del sistema (y del fabricante) para sostener un uso correcto, continuo y contextualizado en condiciones reales (educación, soporte técnico, reemplazos, gestión de fallas, actualizaciones de *software* y vigilancia poscomercialización). En este sentido, Klonoff et al. subrayan que, además de la métrica de exactitud, la toma de decisiones para la adquisición/reembolso debe considerar: a) evidencia publicada en revistas arbitradas; b) exactitud y confiabilidad para la seguridad del paciente, y c) servicios de soporte más allá del sistema en sí, reconociendo que dichos servicios son relevantes para el uso seguro y sostenido. De forma convergente, las GPC ADA 2025 recomiendan que las personas que usan tecnología (incluido MCG) reciban educación y entrenamiento inicial y continuo, y que se evalúen periódicamente la técnica y las habilidades para interpretar y usar datos, precisamente porque los beneficios clínicos y la prevención del daño dependen del uso competente y consistente^{16,36}.

Desde la perspectiva de los estándares iCGM de la FDA, la recomendación cardinal en materia de seguridad es privilegiar sistemas que

implementen controles especiales orientados a minimizar errores clínicamente peligrosos cuando el MCG se integra con otras tecnologías. En el compendio de requisitos resumido por Klonoff et al. (Tabla 6), los controles especiales contemplan umbrales explícitos por rango glucémico (p. ej., en < 70 mg/dl, proporción mínima de lecturas dentro de ± 15 mg/dl; en 70-180 mg/dl y > 180 mg/dl, proporciones dentro de $\pm 15\%$; además, requisitos dentro de $\pm 40/40\%$ y límites para errores de «riesgo extremo», e incluyen también el desempeño de alertas, la seguridad eléctrica/compatibilidad electromagnética y la ciberseguridad. Esta arquitectura regulatoria implica recomendaciones prácticas de seguridad y servicio: a) verificación precomercial con criterios que capturen hipoglucemia e hiperglucemia (donde el error es más dañino); b) validación del funcionamiento de las alarmas (hipoglucemia, hiperglucemia, pérdida de señal), y c) gobernanza de *software* (incluyendo actualizaciones y controles de ciberseguridad) para evitar la degradación del desempeño o vulneraciones de datos. El consenso latinoamericano de Scharf et al. refuerza la idea de que la adhesión a iCGM mejora la seguridad y de que el marco regulatorio debe enfatizar criterios mínimos y la vigilancia pre y poscomercialización, particularmente en mercados donde coexisten dispositivos sin estándares uniformes^{36,38}.

La razón clínica para exigir estos estándares se aprecia en la evidencia agregada: los metaanálisis muestran que cuando el MCG se implementa con protocolos adecuados, se asocia con mejoras clínicamente relevantes en el control glucémico y en parámetros de seguridad (hipoglucemia). En DT2, el metaanálisis de Seidu et al. (ECA) encontró reducción de HbA1c (DM: -0.20% , IC 95%: -0.30 a -0.10), incremento de TIR (DM: $+5.60\%$, IC 95%: 3.40 - 7.80) y reducción de TBR (DM: -0.30 , IC 95%: -0.55 a -0.05); simultáneamente, no observó incremento estadísticamente significativo en eventos adversos (RR: 1.15 , IC 95%: 0.85 - 1.55) ni en eventos adversos relacionados con el dispositivo (RR: 1.00 , IC 95%: 0.71 - 1.42), y reportó mayor satisfacción con el tratamiento (DME: 0.79 , IC 95%: 0.32 - 1.26). En DT1, Alfadli et

Tabla 6. Requisitos de desempeño (*special controls*) para iCGM de la FDA y traducción operativa para evaluación clínica*

Dominio	Rango glucémico/condición	Criterio (zona de acuerdo)	Umbral mínimo y estadístico	Interpretación clínica	Implicación regulatoria/uso
Exactitud puntual (hipoglucemia)	< 70 mg/dl	± 15 mg/dl	Límite inferior unilateral IC 95% > 85%	Exige alta concordancia absoluta en hipoglucemia, donde el riesgo clínico por subestimación/sobreestimación es mayor ^{36,38}	Soporta decisiones no adyuvantes y reduce riesgo en umbrales críticos; requisito clave para interoperabilidad con sistemas conectados ^{36,39}
Exactitud puntual (rango euglucémico)	70-180 mg/dl	$\pm 15\%$	Límite inferior unilateral IC 95% > 70%	Asegura concordancia proporcional en el rango más frecuente de uso clínico, equilibrando desempeño y realismo fisiológico ^{36,38}	Estandariza comparabilidad entre sistemas y facilita implementación en ecosistemas digitales (<i>apps</i> , plumas y bombas inteligentes) ^{36,39}
Exactitud puntual (hiperglucemia)	> 180 mg/dl	$\pm 15\%$	Límite inferior unilateral IC 95% > 80%	Exige desempeño robusto en hiperglucemia, relevante para correcciones de insulina y prevención de cetosis en DM1 ^{36,38}	Requisito para reclamos clínicos consistentes con desempeño real; contribuye a seguridad en automatización de dosis ^{36,39}
Tolerancia ampliada (hipoglucemia)	< 70 mg/dl	± 40 mg/dl	Límite inferior unilateral IC 95% > 98%	Limita discrepancias grandes en hipoglucemia; reduce probabilidad de error clínicamente peligroso ³⁶	Criterio de seguridad «de cola» para evitar fallos catastróficos en umbrales bajos ³⁶
Tolerancia ampliada (rango euglucémico)	70-180 mg/dl	$\pm 40\%$	Límite inferior unilateral IC 95% > 99%	Controla <i>outliers</i> relevantes que podrían inducir decisiones erróneas aun fuera del $\pm 15\%$ ³⁶	Refuerza robustez analítica necesaria para interoperabilidad y uso no adyuvante ³⁶
Tolerancia ampliada (hiperglucemia)	> 180 mg/dl	$\pm 40\%$	Límite inferior unilateral IC 95% > 99%	Limita errores extremos en hiperglucemia, especialmente importantes para correcciones y algoritmos ³⁶	Criterio de seguridad de <i>outliers</i> para integración con sistemas de dosificación automatizada ³⁶
Criterio global de acuerdo	Global (todas las mediciones)	$\pm 20\%$	Límite inferior unilateral IC 95% > 87%	Resumen global que complementa la estratificación por rangos, evitando que el desempeño «promedio» oculte fallas sistemáticas ^{36,38}	Métrica global para comparación y comunicación regulatoria de desempeño mínimo ^{36,38}

(Continúa)

Tabla 6. Requisitos de desempeño (*special controls*) para iCGM de la FDA y traducción operativa para evaluación clínica* (*continuación*)

Dominio	Rango glucémico/ condición	Criterio (zona de acuerdo)	Umbral mínimo y estadístico	Interpretación clínica	Implicación regulatoria/uso
Eventos extremos (discordancia peligrosa)	CGM < 70 mg/dl y referencia > 180 mg/dl	Evento prohibido (0)	Recuento = 0	Evita clasificación errónea crítica (hipo reportada cuando en realidad hay hiper), con alto potencial de daño por intervención inadecuada ³⁶	Criterio de seguridad binario para prevenir errores clínicos inaceptables ³⁶
Eventos extremos (discordancia peligrosa)	CGM > 180 mg/dl y referencia < 70 mg/dl	Evento prohibido (0)	Recuento = 0	Evita clasificación errónea crítica (hiper reportada cuando en realidad hay hipo), con riesgo de sobrecorrección e hipoglucemia grave ³⁶	Criterio de seguridad binario para prevenir errores clínicos inaceptables ³⁶
Exactitud de tendencia (RoC)	CGM RoC > 1 mg/dl/min y ref RoC < -2 mg/dl/min	Discordancia de tendencia	Proporción < 1%	Controla errores direccionales graves durante cambios rápidos, relevantes para alertas predictivas y decisiones terapéuticas ³⁶	Crítico para interoperabilidad (algoritmos) y para desempeño en escenarios dinámicos ³⁶
Exactitud de tendencia (RoC)	CGM RoC < -1 mg/dl/min y ref RoC > 2 mg/dl/min	Discordancia de tendencia	Proporción < 1%	Evita que el sistema indique caída cuando el referente sube rápidamente, escenario de alto riesgo por retraso/lag y decisiones erróneas ³⁶	Refuerza seguridad de integración con dosificación automatizada y sistemas conectados ³⁶

*Los umbrales corresponden a límites inferiores unilaterales del 95% cuando aplica.

FDA: U.S. Food and Drug Administration; IC: intervalo de confianza; iCGM: *integrated continuous glucose monitoring*; LCL: límite inferior (*lower confidence limit*);

Ref: medición de referencia; RoC/ROC: *rate of change* (tasa de cambio).

Adaptada de Klonoff et al.³⁶ y Scharf, et al.³⁸.

al. documentaron beneficios adicionales: HbA1c (DM: -0.37%, IC 95%: -0.52 a -0.21), TIR (+1.68 horas/día, IC 95%: 0.44-2.92) y disminución del tiempo en hipoglucemia (-0.62 horas/día, IC 95%: -1.13 a -0.11), junto con reducción del miedo a la hipoglucemia (DM: -0.30) y aumento de la satisfacción (DM: 0.79)^{33,52-54}.

En los consensos internacionales, la seguridad del MCG se operacionaliza como el uso competente en contexto, lo que se traduce en recomendaciones explícitas de servicio: entrenamiento con equipos experimentados, revisión

de descargas y alineación de la tecnología con las circunstancias del paciente. El consenso internacional sostiene que el uso exitoso del MCG requiere entrenamiento en su aplicación, que los clínicos revisen las descargas y que se considere el contexto vital del paciente. Además, el mismo consenso enfatiza elementos de seguridad ligados a la calibración y a la prevención de la hipoglucemia: la calibración (cuando aplique) debe seguir la guía del fabricante y apoyarse en la glucosa en sangre (SMBG), y se destaca la utilidad de las tasas de cambio (flechas de tendencia)

y de la incorporación de umbrales/alertas para prevenir la hipoglucemia^{16,46}.

NECESIDADES NO CUBIERTAS E INVESTIGACIÓN FUTURA

A pesar de los avances sustantivos en el MCG, persisten necesidades no cubiertas que limitan su impacto clínico, su adopción equitativa y la comparabilidad de la evidencia. En el plano metrológico, sigue existiendo variabilidad entre sistemas y entre generaciones de sensores en condiciones de vida real, particularmente en rangos de hipoglucemia, durante cambios rápidos de glucosa (altas tasas de cambio) y hacia el final del periodo de uso del sensor. Se requieren estudios de desempeño que representen de manera deliberada distribuciones de glucosa y escenarios fisiológicos relevantes (posprandial, ejercicio, enfermedad intercurrente), con comparadores trazables, emparejamiento temporal transparente y un reporte estandarizado que permita comparar estudios y dispositivos.

En seguridad, la fatiga por alarmas, las brechas de disponibilidad de datos (*data gaps*), la comprensión del sensor, los interferentes y la degradación del adhesivo continúan generando fallas funcionales y eventos de uso inseguro. La investigación futura debe priorizar métricas de confiabilidad y de desempeño dinámico (tendencia, retardo/*lag*) con metodologías armonizadas, así como estrategias de mitigación centradas en el usuario (educación, confirmación capilar cuando corresponda y sistemas de soporte). En escenarios hospitalarios y perioperatorios aún falta evidencia robusta sobre la precisión clínica, la integración con flujos de trabajo y los desenlaces duros, lo que condiciona su adopción a gran escala en marcos regulatorios heterogéneos.

Asimismo, la interoperabilidad plantea retos emergentes: ciberseguridad, integridad de datos, trazabilidad de versiones de *software*/algoritmos y control de cambios, especialmente

en ecosistemas conectados (*apps*, plumas inteligentes, sistemas de administración automatizada de insulina). Se requieren estándares de auditoría y de vigilancia poscomercialización que capturen fallos de *software*, problemas de conectividad y sesgos inducidos por las actualizaciones. Finalmente, la brecha de equidad en el acceso y la alfabetización digital sigue siendo crítica. Las futuras líneas de investigación deben incluir evaluaciones de costo-efectividad contextualizadas, la implementación en sistemas de salud con recursos limitados y diseños que incorporen determinantes sociales, para asegurar que el beneficio del MCG sea clínicamente significativo y distribuido de forma justa.

FINANCIAMIENTO

Parte del financiamiento de este proyecto fue otorgado por Laboratorio Abbott.

CONFLICTO DE INTERESES

Los autores declaran no tener conflicto de intereses.

CONSIDERACIONES ÉTICAS

Protección de personas y animales. Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

Confidencialidad, consentimiento informado y aprobación ética. Este consenso se desarrolló mediante revisión de la evidencia y metodología Delphi modificada, sin intervención en pacientes ni acceso a datos personales o clínicos identificables. La participación de los expertos fue voluntaria y la votación individual y anónima, preservando la independencia y transparencia del proceso.

Declaración sobre el uso de inteligencia artificial. Los autores declaran que no se utilizó ningún tipo de inteligencia artificial generativa para la redacción ni la creación de contenido de este manuscrito.

REFERENCIAS

- Ong KL, Stafford LK, McLaughlin SA, Boyko EJ, Vollset SE, Smith AE, et al. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet*. 2023;402(10397):203-34.
- Zhou B, Rayner AW, Gregg EW, Sheffer KE, Carrillo-Larco RM, Bennett JE, et al. Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants. *Lancet*. 2024;404(10467):2077-93.
- Gregory GA, Robinson TIG, Linklater SE, Wang F, Colagiuri S, de Beaufort C, et al. Global incidence, prevalence, and mortality of type 1 diabetes in 2021 with projection to 2040: a modelling study. *Lancet Diabetes Endocrinol*. 2022;10(10):741-60.
- Sweeting A, Hannah W, Backman H, Catalano P, Feghali M, Herman WH, et al. Epidemiology and management of gestational diabetes. *Lancet*. 2024;404(10448):175-92.
- Basto-Abreu AC, López-Olmedo N, Rojas-Martínez R, Aguilar-Salinas CA, De la Cruz-Góngora VW, Rivera-Dommarco J, et al. Prevalence of diabetes and glycemic control in Mexico: national results from 2018 and 2020. *Salud Publica Mex*. 2021;63(6, Nov-Dic):725-33.
- Basto-Abreu A, López-Olmedo N, Rojas-Martínez R, Aguilar-Salinas CA, Moreno-Banda GL, Carnalla M, et al. Prevalencia de prediabetes y diabetes en México: Ensanut 2022. *Salud Publica Mex*. 2023;65:s163-8.
- Wacher NH, Gómez-Díaz RA, Ascencio-Montiel I de J, Rascón-Pacheco RA, Aguilar-Salinas CA, Borja-Aburto VH. Type 1 diabetes incidence in children and adolescents in Mexico: Data from a nation-wide institutional register during 2000-2018. *Diabetes Res Clin Pract*. 2020;159:107949.
- Reyes-Muñoz E, Parra A, Castillo-Mora A, Ortega-González C. Effect of the diagnostic criteria of the International Association of Diabetes and Pregnancy Study Groups on the prevalence of gestational diabetes mellitus in urban Mexican women: a cross-sectional study. *Endocrine Practice*. 2012;18(2):146-51.
- Nava Guerrero EN, Salcedo González A, Hernández Escobar CE, Rivera FC, Mondragón Ortiz EJ, De Alba IM, et al. Prevalence, risk factors and perinatal outcomes of gestational diabetes in Mexican adolescents when applying diagnostic criteria from three different international guidelines. *Int J Diabetes Dev Ctries*. 2021;41(2):212-20.
- Lind M, Polonsky W, Hirsch IB, Heise T, Bolinder J, Dahlqvist S, et al. Continuous glucose monitoring vs conventional therapy for glycaemic control in adults with type 1 diabetes treated with multiple daily insulin injections: the GOLD randomized clinical trial. *JAMA*. 2017;317(4):379-87.
- Beck RW, Riddleworth T, Ruedy K, Ahmann A, Bergenstal R, Haller S, et al. Effect of continuous glucose monitoring on glycaemic control in adults with type 1 diabetes using insulin injections: the DIAMOND randomized clinical trial. *JAMA*. 2017;317(4):371-8.
- Heinemann L, Freckmann G, Ehrmann D, Faber-Heinemann G, Guerra S, Waldenmaier D, et al. Real-time continuous glucose monitoring in adults with type 1 diabetes and impaired hypoglycaemia awareness or severe hypoglycaemia treated with multiple daily insulin injections (HypoDE): a multicentre, randomised controlled trial. *Lancet*. 2018;391(10128):1367-77.
- Martens T, Beck RW, Bailey R, Ruedy KJ, Calhoun P, Peters AL, et al. Effect of continuous glucose monitoring on glycaemic control in patients with type 2 diabetes treated with basal insulin: a randomized clinical trial. *JAMA*. 2021;325(22):2262-72.
- Basu S, Yudkin JS, Kehlenbrink S, Davies JJ, Wild SH, Lipska KJ, et al. Estimation of global insulin use for type 2 diabetes, 2018-30: a microsimulation analysis. *Lancet Diabetes Endocrinol*. 2019;7(1):25-33.
- Gutiérrez JP, Rivera-Dommarco J, Shamah-Levy T, Villalpando-Hernández S, Franco A, Cuevas-Nasu L, et al. Encuesta Nacional de Salud y Nutrición 2012. Resultados Nacionales. Cuernavaca, México: Instituto Nacional de Salud Pública (MX); 2012.
- ElSayed NA, McCoy RG, Aleppo G, Balapattabi K, Beverly EA, Briggs Early K, et al. 7. Diabetes Technology: Standards of Care in Diabetes-2025. *Diabetes Care*. 2025;48(Supplement_1):S146-S166.
- Grunberger G, Sherr J, Allende M, Blevins T, Bode B, Handelsman Y, et al. American Association of Clinical Endocrinology Clinical Practice Guideline: The use of advanced technology in the management of persons with diabetes mellitus. *Endocr Pract*. 2021;27(6):505-37.
- de Bock M, Codner E, Craig ME, Huynh T, Maahs DM, Mahmud FH, et al. Clinical Practice Consensus Guidelines 2022: Glycemic targets and glucose monitoring for children, adolescents, and young people with diabetes. *Pediatr Diabetes*. 2022;23(8):1270-6.
- Tauschmann M, Forlenza G, Hood K, Cardona-Hernandez R, Giani E, Hendriekx C, et al. Clinical Practice Consensus Guidelines 2022: Diabetes technologies: Glucose monitoring. *Pediatr Diabetes*. 2022;23(8):1390-405.
- Secchi-Nicolás NC, Lavalle-González FJ, Garnica-Cuellar JC, Pavia-López AA, Mayorga-Butrón JL, Anda-Garay JC, et al. Guía mexicana de práctica clínica para el diagnóstico y tratamiento en pacientes adultos con diabetes tipo 2. *Rev Mex Endocrinol Metab Nutr*. 2024;10(92).
- Diagnóstico y tratamiento de la diabetes mellitus 2 en la persona adulta mayor. Guía de práctica clínica: Guía de evidencias y recomendaciones [Internet]. México: CENETEC; 2021.
- Pleus S, Eichenlaub M, Dabla PK, Diem P, Eriksson Boija E, Fokkert M, 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*. 2026;580:120728.
- Freckmann G, Pleus S, Eichenlaub M, Eriksson Boija E, Fokkert M, Hinzmann R, 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):1072-81.
- Dávila-Ruales V, Gilón LF, Gómez AM, Muñoz OM, Serrano MN, Henao DC. Evaluating the precision and reliability of real-time continuous glucose monitoring systems in ambulatory settings: a systematic review. *Ther Adv Endocrinol Metab*. 2024;15:20420188241304459.
- Putzu A, Grange E, Schorer R, Schiffer E, Gariani K. Continuous perioperative glucose monitoring in noncardiac surgery. *Eur J Anaesthesiol*. 2025;42(2):162-71.
- Carlier L, De Ponthaud C, Jacqueminet S, Phan F, Gaujoux S, Amouyal C. Perioperative use and accuracy of continuous glucose monitoring: A systematic review. *Diabetes Obes Metab*. 2025;27(10):5393-408.
- Charamba B, Liew A, Nadeem A, Newell J, O'Keeffe DT, O'Brien T, et al. Statistical techniques used in analysing simultaneous continuous glucose monitoring and ambulatory electrocardiography in patients with diabetes: a systematic review. *PLoS One*. 2023;18(2):e0269968.
- Olsen MT, Klarskov CK, Dungu AM, Hansen KB, Pedersen-Bjergaard U, Kristensen PL. Statistical packages and algorithms for the analysis of continuous glucose monitoring data: a systematic review. *J Diabetes Sci Technol*. 2025;19(3):787-809.
- Kumar S, Pallin M, Soldatos G, Teede H. Comparison of continuous glucose monitoring to reference standard oral glucose tolerance test for the detection of dysglycemia in cystic fibrosis: a systematic review. *J Clin Transl Endocrinol*. 2022;30:100305.
- St Clair SL, Ulyatt CM, Corkin MT, Lord LG, Crowther CA, Harding JE, et al. Glucose testing methods: a systematic review and meta-analysis of diagnostic accuracy of point-of-care devices for neonatal hypoglycemia. *J Pediatr*. 2025;278:114438.
- Laha S, Rajput A, Laha SS, Jadhav R. A concise and systematic review on non-invasive glucose monitoring for potential diabetes management. *Biosensors (Basel)*. 2022;12(11):965.
- Eichenlaub M, Pleus S, Rothenbühler M, Bailey TS, Bally L, Brazg R, et al. Comparator data characteristics and testing procedures for the clinical performance evaluation of continuous glucose monitoring systems. *Diabetes Technol Ther*. 2024;26(4):263-75.
- Alfadli SF, Alotaibi YS, Aqdi MJ, Almozan LA, Alzubaidi ZB, Altemani HA, et al. Effectiveness of continuous glucose monitoring systems on glycaemic control in adults with type 1 diabetes: A systematic review and meta-analysis. *Metabol Open*. 2025;27:100382.
- Zhou Y, Sardana D, Kuroko S, Haszard JJ, de Block MI, Weng J, et al. Comparing the glycaemic outcomes between real-time continuous glucose monitoring (rt-CGM) and intermittently scanned continuous glucose monitoring (isCGM) among adults and children with type 1 diabetes: a systematic review and meta-analysis of randomized controlled trials. *Diabet Med*. 2024;41(3).

35. Zeng B, Gao L, Yang Q, Jia H, Sun F. Automated insulin delivery systems in children and adolescents with type 1 diabetes: a systematic review and meta-analysis of outpatient randomized controlled trials. *Diabetes Care*. 2023;46(12):2300-7.
36. Klonoff DC, Gabbay M, Moon SJ, Wilmot EG. Importance of FDA-integrated continuous glucose monitors to ensure accuracy of continuous glucose monitoring. *J Diabetes Sci Technol*. 2025;19(5):1392-9.
37. Battelino T, Alexander CM, Amiel SA, Arreaza-Rubin G, Beck RW, Bergenstal RM, et al. Continuous glucose monitoring and metrics for clinical trials: an international consensus statement. *Lancet Diabetes Endocrinol*. 2023;11(1):42-57.
38. Scharf M, Feriz K, Yepes C, Contreras A, Lequi L, Sanhueza D, et al. Consensus statement on standardizing CGM evaluation metrics in Latin America: an expert approach. *Diabetol Metab Syndr*. 2025;17(1):291.
39. Krouwer JS. Minimum requirements for continuous glucose monitors: What does it mean? *J Diabetes Sci Technol*. 2025;19(5):1417-8.
40. Chen D, Lin B, Liu Z, Lei M, Yang Y, Yao B, et al. Effect of real-time continuous glucose monitoring vs. flash glucose monitoring on glycemic control in adults with type 1 diabetes mellitus: a systematic review and meta-analysis. *Metab Syndr Relat Disord*. 2024;22(10):709-16.
41. Richardson KM, Jospe MR, Bohlen LC, Crawshaw J, Saleh AA, Schembre SM. The efficacy of using continuous glucose monitoring as a behaviour change tool in populations with and without diabetes: a systematic review and meta-analysis of randomised controlled trials. *Int J Behav Nutr Phys Act*. 2024;21(1):145.
42. Uhl S, Choure A, Rouse B, Loblack A, Reaven P. Effectiveness of continuous glucose monitoring on metrics of glycemic control in type 2 diabetes mellitus: a systematic review and meta-analysis of randomized controlled trials. *J Clin Endocrinol Metab*. 2024;109(4):1119-31.
43. Aronson R, Abitbol A, Bajaj HS, Cheng AYY, Christopoulos S, Harris SB, et al. Continuous glucose monitoring in noninsulin-treated type 2 diabetes: A critical review of reported trials with an updated systematic review and meta-analysis of randomised controlled trials. *Diabetes Obes Metab*. 2025;27(11):6220-42.
44. Di Molfetta S, Caruso I, Cignarelli A, Natalicchio A, Perrini S, Laviola L, et al. Professional continuous glucose monitoring in patients with diabetes mellitus: a systematic review and meta-analysis. *Diabetes Obes Metab*. 2023;25(5):1301-10.
45. Di Molfetta S, Di Gioia L, Caruso I, Cignarelli A, Green SC, Natale P, et al. Efficacy and safety of different hybrid closed loop systems for automated insulin delivery in people with type 1 diabetes: a systematic review and network meta-analysis. *Diabetes Metab Res Rev*. 2024;40(6):e3842.
46. Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International consensus on use of continuous glucose monitoring. *Diabetes Care*. 2017;40(12):1631-40.
47. Podwojniak A, Flemming J, Tan IJ, Ghani H, Neubauer Z, Jones A. Cutaneous adverse effects from diabetes devices in pediatric patients with type 1 diabetes mellitus: systematic review. *JMIR Dermatol*. 2024;7:e59824.
48. Natale P, Chen S, Chow CK, Cheung NW, Martinez-Martin D, Caillaud C, et al. Patient experiences of continuous glucose monitoring and sensor-augmented insulin pump therapy for diabetes: a systematic review of qualitative studies. *J Diabetes*. 2023;15(12):1048-69.
49. Barchiesi MA, Calabrese A, Costa R, Di Pillo F, D'Uffizi A, Tiburzi L, et al. Continuous glucose monitoring in type 2 diabetes: a systematic review of barriers and opportunities for care improvement. *Int J Qual Health Care*. 2025;37(3):mzaf046.
50. Obermaier K, Schmelzeisen-Redeker G, Schoemaker M, Klötzer HM, Kirchsteiger H, Eikmeier H, et al. Performance evaluations of continuous glucose monitoring systems: precision absolute relative deviation is part of the assessment. *J Diabetes Sci Technol*. 2013;7(4):824-32.
51. Lu J, Ying Z, Wang P, Fu M, Han C, Zhang M. Effects of continuous glucose monitoring on glycaemic control in type 2 diabetes: a systematic review and network meta-analysis of randomized controlled trials. *Diabetes Obes Metab*. 2024;26(1):362-72.
52. Seidu S, Kunutsor SK, Ajjan RA, Choudhary P. Efficacy and safety of continuous glucose monitoring and intermittently scanned continuous glucose monitoring in patients with type 2 diabetes: a systematic review and meta-analysis of interventional evidence. *Diabetes Care*. 2024;47(1):169-79.
53. Cavalcante Lima Chagas G, Teixeira L, Clemente MRC, Cavalcante Lima Chagas R, Santinelli Pestana DV, Rodrigues Silva Sombra L, et al. Use of continuous glucose monitoring and point-of-care glucose testing in hospitalized patients with diabetes mellitus in non-intensive care unit settings: a systematic review and meta-analysis of randomized controlled trials. *Diabetes Res Clin Pract*. 2025;220:111986.
54. Patel PM, Thomas D, Liu Z, Aldrich-Renner S, Clemons M, Patel BV. Systematic review of disparities in continuous glucose monitoring and insulin pump utilization in the United States: key themes and evidentiary gaps. *Diabetes Obes Metab*. 2024;26(10):4293-301.
55. Mayorga J, Velasco L, Ochoa F. Guías de práctica clínica basadas en evidencia, cerrando la brecha entre el conocimiento científico y la toma de decisiones clínicas. Documento de la serie MBE, 3 de 3. *Gac Mex Oncol*. 2015;14(6):329-34.