Evaluating the Cost-Utility of Continuous Glucose Monitoring in Individuals with Type 1 Diabetes: A Systematic Review of the Methods and Quality of Studies Using Decision Models or Empirical Data

Lisa A. de Jong[1] · Xinyu Li[2] · Sajad Emamipour[3] · Sjoukje van der Werf[4] · Maarten J. Postma[1,5,6,7] · Peter R. van Dijk[8,10] · Talitha L. Feenstra[2,9]

Abstract

Introduction This review presents a critical appraisal of differences in the methodologies and quality of model-based and empirical data-based cost-utility studies on continuous glucose monitoring (CGM) in type 1 diabetes (T1D) populations. It identifies key limitations and challenges in health economic evaluations on CGM and opportunities for their improvement.

Methods The review and its documentation adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for systematic reviews. Searches for articles published between January 2000 and January 2023 were conducted using the MEDLINE, Embase, Web of Science, Cochrane Library, and Econlit databases. Published studies using models and empirical data to evaluate the cost utility of all CGM devices used by T1D patients were included in the search. Two authors independently extracted data on interventions, populations, model settings (e.g., perspectives and time horizons), model types and structures, clinical outcomes used to populate the model, validation, and uncertainty analyses. They subsequently met to confirm consensus. Quality was assessed using the Philips checklist for model-based studies and the Consensus Health Economic Criteria (CHEC) checklist for empirical studies. Model validation was assessed using the Assessment of the Validation Status of Health-Economic decision models (AdViSHE) checklist. The extracted data were used to generate summary tables and figures. The study protocol is registered with PROSPERO (CRD42023391284).

Results In total, 34 studies satisfied the selection criteria, two of which only used empirical data. The remaining 32 studies applied 10 different models, with a substantial majority adopting the CORE Diabetes Model. Model-based studies often lacked transparency, as their assumptions regarding the extrapolation of treatment effects beyond available evidence from clinical studies and the selection and processing of the input data were not explicitly stated. Initial scores for disagreements concerning checklists were relatively high, especially for the Philips checklist. Following their resolution, overall quality scores were moderate at 56%, whereas model validation scores were mixed. Strikingly, costing approaches differed widely across studies, resulting in little consistency in the elements included in intervention costs.

Discussion and Conclusion The overall quality of studies evaluating CGM was moderate. Potential areas of improvement include developing systematic approaches for data selection, improving uncertainty analyses, clearer reporting, and explaining choices for particular modeling approaches. Few studies provided the assurance that all relevant and feasible options had been compared, which is required by decision makers, especially for rapidly evolving technologies such as CGM and insulin administration. High scores for disagreements indicated that several checklists contained questions that were difficult to interpret consistently for quality assessment. Therefore, simpler but comprehensive quality checklists may be needed for model-based health economic evaluation studies.

Extended author information available on the last page of the article

Key Points for Decision Makers

Cost-utility studies inform decision makers about the value for money of CGM devices, which also requires clarity on the combinations with different insulin pumps or injections. Our review findings showed that the overall quality of current economic evaluations in this field is moderate.

Referring to existing guidelines, studies could be improved by adopting more systematic approaches to selection of input data, by refining uncertainty analyses, and by clearly reporting methods used. Importantly, researchers should better explain their choices in the modeling process.

1 Introduction

Accurate and timely monitoring of glucose levels is critical for individuals to manage diabetes effectively and achieve optimal blood glucose control. Monitoring glucose levels empowers patients and their caregivers to fine-tune insulin dosages and make lifestyle and dietary changes. Ultimately, optimal glycemic control can postpone or even prevent the onset and progression of micro- and macrovascular complications [1, 2]. Historically, individuals with diabetes have self-monitored their capillary blood glucose levels using finger pricks, which has prevailed since the late 1970s [3]. This method causes discomfort and is time-consuming, discouraging many patients from applying it frequently throughout the day. Moreover, it yields isolated glucose readings and does not offer insights into trends in glucose levels.

Continuous glucose monitoring (CGM) devices, i.e., glucose sensors, may effectively overcome these barriers. Different types of glucose monitoring devices exist, namely intermittently scanned (is-) and real-time (rt-) CGM. Users adopting is-CGM devices must proactively scan (or 'flash') the sensor using a reader or smartphone to obtain a glucose measurement. By contrast, rt-CGM systems provide a continuous data flow. Both systems offer the option of setting alarms at predetermined glucose levels that warn users in case of imminent hypo- or hyperglycemia. CGM devices can be used in conjunction with various insulin administration methods, including insulin injections and pumps. While is-CGM can be used alongside an insulin pump, direct connectivity between the two is not possible. The use of an rt-CGM device in combination with certain insulin pumps allows for a 'closed-loop' system. Over the last decade, closed-loop systems have become increasingly sophisticated. While they initially offered low-glucose suspension in case of hypoglycemia, they now incorporate features enabling anticipation of both hypoglycemia and hyperglycemia. CGM devices can be used by all individuals with diabetes for whom insulin is the primary treatment, which mainly concerns persons with type 1 diabetes (T1D) [8].

Studies have shown that CGM devices improve patients' glucose management and quality of life (QoL), decrease the disease burden, and reduce diabetes-related hospital admissions of individuals with T1D [4, 5]. While several studies have demonstrated these beneficial effects, another aspect of decision making relates to their cost effectiveness [6, 7]. Multiple cost-effectiveness studies conducted for different CGM devices have been published since their entry into the market. Given rapid advances in technology and varying coverage in different countries, the choices of interventions, comparators, and target populations reported in these studies differ.

A recently published systematic review [8] on the cost effectiveness of various types of CGM in individuals with T1D aimed to summarize and clarify study results to provide guidance on the appropriateness of implementing CGM technology based on whether the result is cost effective or not. Its findings indicated that CGM appears to be a cost-effective tool for individuals with T1D, especially in the subgroup of individuals with suboptimal management or those at risk of severe hypoglycemia. However, incremental cost-effectiveness ratios (ICERs) and quality-adjusted life-years (QALY) gained varied widely between studies (US$18,734–US$99,941 per QALY gained and 0.76–2.99 QALYs gained, respectively).

These disparities may be partly linked to differences in the devices used, given the rapid development of CGM systems in recent years. Another important reason for differences could relate to ways of translating the benefits of better glucose monitoring and more controlled insulin administration into lifetime QALY gains. However, these reasons were not further analyzed in the above review, which presented a simple overview of the studies and their outcomes. Further investigations are thus needed to explain differences in the studies' methodologies, quality, and validity, and their potential impact on the results. Establishing state-of-the-art economic evaluations on CGM is important to raise the bar for future research, while also enabling reviewers (e.g., policy makers) to gain a better understanding of the pros and cons of existing studies and improve study designs and reporting.

Therefore, the primary objective of this review is to critically appraise differences in the methodologies, quality, and

validity of studies based on models and empirical data. We examine methods and quality of cost-utility analyses (CUAs) rather than compare quantitative results to explain some of the variations in outcomes and to identify key limitations, challenges, and opportunities that need to be addressed in future CUAs of CGM for individuals with T1D. Given rapid developments in this field, our secondary objectives were to provide recommendations that would benefit future cost-effectiveness studies on CGM and to help decision makers in diverse settings to identify studies that are relevant to them without needing to delve into each study.

2 Methods

2.1 Protocol and Registration

A protocol was developed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses-Protocol (PRISMA-P) checklist [9, 10]. The protocol was registered in PROSPERO (registration number CRD42023391284) [11] and subsequently published [12]. The conduct and reporting of this review adhered to this protocol, following the updated PRISMA 2020 guidelines [13].

2.2 Search Strategy

We searched the following electronic databases: MEDLINE (PubMed), Embase (embase.com), Web of Science (Clarivate), the Cochrane Central Register of Controlled Trials (CENTRAL), and Econlit (EBSCO). Additionally, we used the Google Scholar search engine and searched cited references in previous systematic literature reviews on this topic. The reference lists of identified papers were also screened for additional articles.

The search strategy centered on three terms: (1) CGM, (2) cost-utility, and (3) T1D. The search was limited to English-language articles published between 2000 and 2023, as the US FDA approved the first CGM device in June 1999 [14]. Details can be found in Online Resource 1 in the Supplementary Information.

2.3 Eligibility Criteria

The Population, Intervention, Comparator, Outcome and Study (PICOS) design was defined as follows: P: T1D; I: any CGM device; C: any; O: ICER as cost per QALY; and S: model- and empirical data-based cost-effectiveness studies. The review protocol provides details of the PICOS [12]. CGM was broadly defined as any system, including both is-CGM and rt-CGM, capable of monitoring blood glucose concentrations on a 'continuous basis' [15, 16]. No selection was made based on insulin administration methods

combined with CGM [16, 17]. Studies on bi-hormonal closed loops (e.g., an artificial pancreas) and do-it-yourself hybrid closed loop (HCL) devices were excluded, as these interventions are still under development. We included cost-utility studies based on models and empirical data (i.e., trial or observational data). However, we excluded studies that reported results only in terms of cost per life-year gained or other clinical outcomes, or those that reported only costs or clinical outcomes. This is because a focus on life-years gained would lead to the underestimation of expected health benefits, and it is difficult to compare clinical outcomes across studies. Studies conducted on individuals diagnosed with T1D were included irrespective of age, ethnicity, comorbidities, and background therapies administered. Studies considering both T1D and T2D were included only if the results for each population were described separately, as their results and inputs would likely differ.

2.4 Study Process

The search results were imported into EndNote 20, and a structured method was used to remove duplicates [17]. Titles and abstracts were then exported to the review manager Rayyan for screening purposes [18]. Two reviewers who independently reviewed the titles and abstracts assessed all the studies against the eligibility criteria. Full-text screening of the remaining articles was performed independently by the same two reviewers using predefined selection criteria based on PICOS. Inconsistencies and disagreements among the researchers were resolved through discussion.

2.5 Data Extraction Procedures

Data extraction was performed in Microsoft Excel using a prespecified data extraction table (Online Resource 2). Extracted data comprised basic information on the studies (setting, population, intervention, comparator, perspective, and year of publication); the model structure or empirical study design employed; diabetes-related complications considered; methods of modeling the effectiveness of CGM and data sources for effectiveness; input data used for costs and QoL; handling of uncertainty; model validity; and quality assessment. Additionally, incremental costs, QALY and ICER data, as well as conflict of interest statements and study funding data were collected. The studies were allocated to four researchers, who performed data extraction in such a way that each study was examined independently by two authors. After the first five studies had been extracted, the researchers met to discuss their experiences with the data extraction form. In light of this discussion, the form was improved, and redundant columns, notably the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist [19], were removed, as items on this list

were already covered in other sections. Explanatory notes were also added. Finally, after all the data extraction forms had been completed, the authors met in pairs to compare results and confirm consensus. During meetings with all the researchers, pending issues were discussed and the final results were generated.

2.6 Quality and Validation Assessment

Following the recommendations of Wijnen et al. [20], a quality assessment of the cost-utility studies was performed using the checklist developed by Philips et al. [21] for model-based evaluations, and the Consensus Health Economic Criteria (CHEC)-extended checklist [22, 23] for economic evaluations based on trial and observational data. The Assessment of the Validation Status of Health-Economic decision models (AdViSHE) checklist [24] was used for validation. We discussed each of these checklists critically to develop a scoring system for the quality assessment. Two reviewers independently assigned a score to each study in light of the information contained in the publication or associated background papers. While we considered any completed checklists that were provided as part of the included study itself, we completed the checklist for each study ourselves to ensure that the responses and scores remained consistent across studies. The Philips checklist has 57 sub-questions (Online Resource 3). Some of the items on this checklist were scored at the model level, thereby avoiding repetitive data extraction for studies using the same model and citing the same background references. However, this was only possible for items on the modeling methods used; the responses to other items, entailing setting-specific considerations were scored at the study level. The CHEC-extended checklist comprises 19 questions requiring 'yes' or 'no' responses. Responses to items in both checklists were 'yes', 'no', 'not reported', or 'not applicable'. Additional explanations were recorded in separate columns. Completion of the Philips checklist required several additional meetings to achieve consensus regarding data extraction among the co-authors. The AdViSHE checklist comprised 13 items covering different aspects of the validation of the conceptual model, the input data, the implemented software program, and model outcomes (Online Resource 2) [24]. We have only considered the model validation items that were reported in the study or its background literature, i.e., no additional questions were sent to the authors of the studies regarding activities performed but not reported.

2.7 Data Synthesis

We synthesized the extracted data qualitatively, focusing on the methods and models applied. Tables and graphs were constructed depicting the general characteristics of the studies, the types of interventions and comparators, sources for obtaining baseline characteristics and CGM effectiveness, input data, model structure, methodological quality, and model validation. The extracted incremental ICERs were converted at US dollar values in 2022 by indexing the ICERs according to the price year in the specific country's currency and then converting this value into US dollars using purchasing power parities obtained from Organisation for Economic Co-operation and Development (OECD) data [25–27]. The range of ICERs was reported, and each individual result was plotted along a cost-effectiveness plane to indicate the distribution of outcomes in terms of incremental costs and QALYs. The various comparisons included in the studies and the results of the quality and validation assessments were summarized and presented in stacked bar charts.

3 Results

Out of a total of 1736 screened articles, 34 cost-utility studies on CGM for individuals with T1D that were based on models or empirical data were eligible for inclusion in this review. Figure 1 depicts the process of selecting studies, and Table 1 presents an overview of each of the included studies, sorted by the type of analysis (based on empirical data or models), author(s), and the publication year. A wide range of ICERs was reported, varying from dominant (cost saving with QALY gains) to $4,858,767 per QALY. Most of the reported ICERs were either below $50,000 per QALY (62%) or below $100,000 per QALY (75%) [Fig. 2]. Estimated incremental QALYs ranged from − 0.01 to 3.81 QALYs per patient, and estimated incremental costs per patient ranged from − $22,853 to $224,678 (Online Resource 4).

The studies, which were conducted in 14 countries, entailed multiple comparisons in terms of CGM device and insulin administration used (Table 1). Figure 3 presents a summary of the comparative analyses (details are provided in Online Resource 5). Transparency regarding the type of devices evaluated was low for both the method of insulin administration and CGM, indicated by at least one disagreement among reviewers on methods used in 21/34 studies (62%). A further indicator of non-transparency was the absence of clear definitions of interventions and comparators in the Objectives or Methods sections in many of the studies. These had to be deduced from the clinical data or cost tables. Three studies evaluated CGM devices without specifying the method of insulin administration [30–32]; one of these studies did not even specify the CGM type (although costs were included for three different rt-CGM devices) [32]. The specific brand and type of pump or sensor was reported in only 24 of the 34 studies. Several studies that evaluated integrated systems (sensor-augmented pump [SAP] therapy and HCL) reported the insulin pump device but not the specific sensor used [33–35]. One study's comparison of an

integrated system (SAP) to the same system without connectivity between the insulin pump and rt-CGM device is questionable from a clinical perspective [36].

Some studies compared various CGM modes [30, 33, 36–38], yielding a total of 40 different comparisons. The most common comparison (in 15/40 studies) involved the SAP (rt-CGM+SAP) and self-monitoring of blood glucose (SMBG) combined with either continuous subcutaneous insulin infusion (CSII) [34–36, 39–46], multiple daily injections (MDI) [36, 47, 48], or a combination of both [37]. One study compared a combination of rt-CGM and MDI with SMBG and MDI, and is-CGM with MDI [38]. Two studies compared rt-CGM with SMBG but did not specify the method of insulin administration [30, 32]. Four studies compared is-CGM with SMBG and MDI [7, 31, 49]. Eight studies compared HCL systems with SMBG [50–53], is-CGM [33, 54, 55], or with SAP [33] in combination with varying insulin administration methods. Thus, relatively few studies compared HCL systems with SAP, or SAP with rt-CGM plus CSII or MDI, which offer more advanced monitoring and insulin delivery options. The comparator frequently involved SMBG.

Most of the studies were model-based CUAs, with two exceptions: a trial-based study [39] and one based on observational data [7]. Two other studies included a within-trial analysis and a model-based analysis with a lifetime horizon [28, 56]. Of the 32 model-based studies, 24 (75%) applied the CORE Diabetes Model, and one study applied the Sheffield Type 1 Diabetes Model [56]. The remaining 7 models were newly developed, although 4 were variants of each other [30, 32, 37, 57]. The majority of the studies were conducted from a societal perspective (15/34), including all costs that were substantially affected by the intervention within and beyond the healthcare sector. Studies adopting a healthcare system perspective (11/34) and payer's perspective (8/34) [Table 1] ranked second and third, respectively. The four studies based on empirical data used a within-study time horizon of 6 or 12 months. They found that the CGM variant they investigated was cost effective or even cost saving. All except two model-based studies used a lifetime horizon. Only four model-based studies (all publicly sponsored) found that CGM variants were not cost effective according to their reported willingness-to-pay (WTP) thresholds [36, 37, 47, 53]. Considering a WTP threshold of $100,000 per QALY, 12 interventions (Fig. 2)

Fig. 1 Process of selecting studies following the PRISMA 2020 guidelines. *PRISMA* Preferred Reporting Items for Systematic Reviews and Meta-Analyses
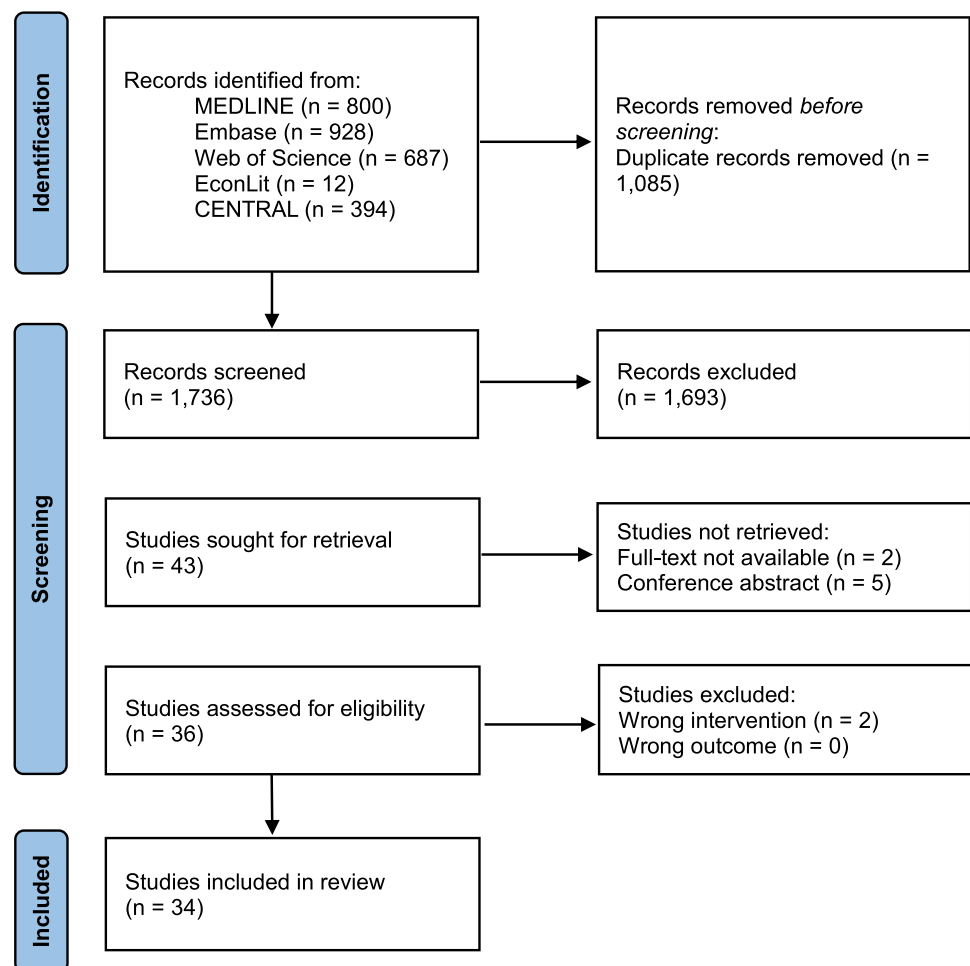
Table 1 Overview of cost-utility studies included in the review

Publication (first author, year, country)	Based on a model or empirical data/version	Patient population	Intervention(s)	Comparator(s)	Perspective[a]	Time horizon	Cost effective?[b]
Emamipour (2022), The Netherlands [7]	Observational study	Adults with T1D	is-CGM + MDI	NR + NR	Societal	12 months	Yes
Huang (2010), USA [28]	Patient-level Markov model	Individuals with T1D with CSII or MDI of insulin, with HbA1c levels ≤10.0%	rt-CGM + MDI or CSII	SMBG + MDI or CSII	Societal	6 months; lifetime	Yes
Ly (2014), Australia [39]	Trial-based	Children, adolescents, and adults with T1D	rt-CGM + SAP	SMBG + CSII	Healthcare system	6 months	Yes
Wan (2018), US [56]	Trial-based and model-based (Sheffield type 1 Diabetes Model)	Adults with T1D whose HbA1c levels were elevated with MDI of insulin	rt-CGM + MDI	SMBG + MDI	Societal	6 months; lifetime	Yes
Bilir (2018), Sweden [49]	CORE Diabetes Model (v. 9.0)	Adults with well-controlled T1D (HbA1c ≤7.5%) with MDI of insulin or CSII therapy and testing glucose levels at least 10 times/week	is-CGM + MDI or CSII	SMBG + MDI or CSII	Payer	Lifetime	Yes
Chaugule (2017), Canada [58]	CORE Diabetes Model (v. 9.0)	Adults with T1D whose HbA1c levels were elevated with MDI of insulin	rt-CGM + MDI	SMBG + MDI	Healthcare system	Lifetime	Yes
Conget (2018), Spain [34]	CORE Diabetes Model	Individuals with T1D at a high risk for hypoglycemia	rt-CGM + SAP	SMBG + CSII	Societal	Lifetime	Yes
Gómez (2016), Colombia [48]	CORE Diabetes Model (v. 8.5)	Individuals with T1D aged >11 years	rt-CGM + SAP	SMBG + MDI	Payer	Lifetime	Yes
Isitt (2022), Australia [38]	CORE Diabetes Model	Adults with T1D whose HbA1c levels were elevated with MDI of insulin	rt-CGM + MDI	SMBG and is-CGM + MDI	Healthcare system	Lifetime	Yes
Jendle (2017), Sweden [35]	CORE Diabetes Model	Individuals with T1D: (1) with an increased risk for hypoglycemia; (2) with uncontrolled HbA1c at baseline	rt-CGM + SAP	SMBG + CSII	Societal	Lifetime	Yes
Jendle (2019), Sweden [51]	CORE Diabetes Model	Individuals with T1D aged 14–75 years who had been on CSII therapy with or without CGM for >6 months	rt-CGM in HCL	SMBG + CSII	Societal	Lifetime	Yes
Jendle (2021), Sweden [54]	CORE Diabetes Model (v. 9.0)	Individuals with T1D for >3 months	rt-CGM in HCL	is-CGM + MDI or CSII	Societal	Lifetime	Yes
Kamble (2012), US [47]	CORE Diabetes Model (v. 7.0)	Adults with inadequately controlled T1D	rt-CGM + SAP	SMBG + MDI	Healthcare system	Lifetime	No
Lambadiari (2022), Greece [33]	CORE Diabetes Model	Individuals with T1D	rt-CGM in HCL	rt-CGM + SAP or is-CGM + MDI	Societal	Lifetime	Yes

Table 1 (continued)

Publication (first author, year, country)	Based on a model or empirical data/version	Patient population	Intervention(s)	Comparator(s)	Perspective[a]	Time horizon	Cost effective?[b]
Nicolucci (2018), Italy [46]	CORE Diabetes Model	Individuals with T1D: (1) at increased risk for hypoglycemia; (2) with uncontrolled HbA1c	rt-CGM + SAP	SMBG + CSII	Societal	Lifetime	Yes
Riemsma (2016), UK [36]	CORE Diabetes Model (v. 8.5)	Individuals with T1D who were eligible to use an insulin pump	rt-CGM + SAP	rt-CGM or SMBG + CSII or MDI	Healthcare system	Lifetime	No
Roze (2015), Sweden [40]	CORE Diabetes Model	Individuals with T1D	rt-CGM + SAP	SMBG + CSII	Societal	Lifetime	Yes
Roze (2016), France [41]	CORE Diabetes Model	Individuals with T1D: (1) with an impaired awareness of hyperglycemia and therefore an elevated risk for it; (2) with uncontrolled HbA1c	rt-CGM + SAP	SMBG + CSII	Payer	Lifetime	Yes
Roze (2016), UK [42]	CORE Diabetes Model	Individuals with T1D with poor glycemic control	rt-CGM + SAP	SMBG + CSII	Healthcare system	Lifetime	Yes
Roze (2017), Denmark [44]	CORE Diabetes Model	Individuals with T1D: (1) with hyperglycemia at baseline; (2) with an increased risk for hypoglycemia	rt-CGM + SAP	SMBG + CSII	Societal	Lifetime	Yes
Roze (2019), The Netherlands [45]	CORE Diabetes Model	Individuals with T1D: (1) with hyperglycemia at baseline; (2) with an increased risk for hypoglycemia	rt-CGM + SAP	SMBG + CSII	Societal	Lifetime	Yes
Roze (2019), Turkey [43]	CORE Diabetes Model	Individuals with T1D: (1) with poor glycemic control at baseline; (2) at increased risk for hypoglycemia	rt-CGM + SAP	SMBG + CSII	Payer	Lifetime	Yes
Roze (2020), UK [59]	CORE Diabetes Model	Individuals with T1D: (1) reflecting the DIAMOND trial T1D population; (2) reflecting the DIAMOND trial T1D population with a baseline HbA1c of at least 8.5% (69 mmol/mol)	rt-CGM + MDI	SMBG + MDI	Payer	Lifetime	Yes
Roze (2021), Canada [60]	CORE Diabetes Model	Adults with T1D	rt-CGM + MDI	SMBG + MDI	Payer	Lifetime	Yes
Roze (2021), UK [50]	CORE Diabetes Model	Adults and adolescents with T1D	rt-CGM in HCL	SMBG + CSII	Payer	Lifetime	Yes
Roze (2021), France [61]	CORE Diabetes Model (v. 9.0)	Adults with T1D	rt-CGM + MDI	SMBG + MDI	Payer	Lifetime	Yes

Table 1 (continued)

Publication (first author, year, country)	Based on a model or empirical data/version	Patient population	Intervention(s)	Comparator(s)	Perspective[a]	Time horizon	Cost effective?[b]
Serné (2022), The Netherlands [55]	CORE Diabetes Model	Individuals with T1D	rt-CGM in HCL	is-CGM + MDI or CSII	Societal	Lifetime	Yes
Zhao (2021), China [31]	CORE Diabetes Model (v. 9.5)	Individuals with T1D	is-CGM + NR	SMBG + NR	Societal	Lifetime	Yes
Garcia-Lorenzo (2018), Spain [57]	Markov model adapted from McQueen et al. [32]	Individuals with T1D without complications at the baseline	rt-CGM + MDI or CSII	SMBG + MDI or CSII	Healthcare system	Lifetime	No
Health Quality Ontario (2018), Canada [37]	A state-transition model structure adapted from McQueen et al. [32]	Individuals with T1D	SAP or rt-CGM + MDI or CSII	SMBG + MDI or CSII	Healthcare system	Lifetime	No
McQueen (2011), USA [32]	Study's own Markov model with input from the CDC Cost-Effectiveness Group model	Adults with T1D	rt-CGM + NR	SMBG + NR	Societal	Lifetime	Yes
Pease (2020), Australia [52]	Study's own Markov model	Adults with T1D	rt-CGM in HCL	SMBG + MDI	Healthcare system	Lifetime	Yes
Pease (2022), Australia [53]	Patient-level Markov model	Young people with T1D	rt-CGM in HCL	SMBG + MDI or CSII	Healthcare system	13 years	Yes
Rotondi (2022), Canada [30]	A Markov model adapted from an Ontario Health (OH) [37] report, Garcia-Lorenzo et al. [57] and McQueen et al. [32]	Adults with T1D aged 18–64 years	rt-CGM or is-CGM + NR	SMBG + NR	Healthcare system	20 years	Yes

CSII continuous subcutaneous insulin infusion, *HCL* hybrid closed loop, *is-CGM* intermittently-scanned continuous glucose monitoring, *MDI* multiple daily injections, *NR* not reported, *rt-CGM* real-time continuous glucose monitoring, *SAP* sensor-augmented pump, *SMBG* self-monitoring blood glucose, *T1D* type 1 diabetes

[a]In the base-case analysis of the economic evaluation

[b]Authors' conclusion

could not be considered cost effective [28, 36, 37, 47, 56, 57]. All 25 studies that were sponsored by the device manufacturer found the CGM devices to be cost effective or cost saving (Online Resource 6).

3.1 Economic Evaluation Methodology

3.1.1 Studies Based on Empirical Data

One study was based on a prospective nationwide Dutch observational study [7] and the remaining three studies were based on unblinded, multicenter, randomized controlled trials (RCTs) [28, 39, 56] (Online Resource 7). The studies investigated 95–381 individuals with T1D and were conducted over 6 or 12 months between late 2009 and 2016. The observational study only collected QoL data (EuroQol questionnaire EQ-5D-3L), with costs and healthcare resource use derived from linked health insurance data. The three RCTs collected cost data as well as time trade-off [28], EQ-5D-3L [39] or EQ-5D-5L [56] data from questionnaires.

3.1.2 Model-Based Studies

Most of the model-based studies (*n* = 24) used the CORE Diabetes Model (Online Resource 8). CORE is an individual patient-level Markov model used for cost-effective analyses in T1D and T2D. Two publications by Palmer et al. [62, 63] first introduced the CORE Diabetes Model and its validation. Most, although not all of the reviewed CUAs referenced these two papers, usually alongside a 2014 validation paper by McEwan et al. [64]. Although CORE is commonly used to assess the cost utility of diabetes interventions, most studies applying it did not provide a rationale for their choice of model. One other study chose the Sheffield T1D policy model [65], which is an individual patient-level simulation model, explicitly arguing that this model was designed and validated solely using T1D studies and trials [56]. They provided additional justification, noting that HbA1c was included in most of the risk equations, the model was validated against major T1D trials, and was completely transparent. This was the only study that explicitly discussed the model choice. Four other studies used self-developed Markov state-transition models, which were all slightly differing variations of the model developed by McQueen et al. [32]. The remaining three studies by Huang et al. and Pease et al. [28, 52, 53] used models developed by the authors themselves. Two were cohort-based Markov models and the third was a patient-level Markov model. It is unclear why different model structures were used for the two studies by Pease et al., as both studies evaluated the cost-utility of HCL: one focused on adults [52] and the other on youth [53].

Although different model structures were used, all the models applied a 1-year cycle, while allowing for the occurrence of more than one hypoglycemic event annually by modeling this into the background or using a shorter cycle (e.g., 3 months in the CORE Diabetes Model). All but three studies [28, 32, 57] considered hypoglycemic events. One of these studies motivated the exclusion of hypoglycemic events by lack of effect on such events in the clinical study on which the effect size was based [57]. However, most studies did not provide a clear definition for hypoglycemia or the distinction between severe and non-severe hypoglycemic events.

Varying amounts of diabetes complications were included in the models, even in studies applying the same model (Online Resource 9). The inclusion or exclusion of specific complications was rarely justified but may have been due to differing versions of the model used or the exclusion of certain complications in the event of an absence of local data for estimating costs. For most studies, the included complications were apparent, although often they had to be inferred from tables on costs and results. For one of the studies applying the CORE Diabetes Model, information on complications was missing entirely, although it can be assumed that they would likely be similar to those considered in most studies applying this model [49]. Macrovascular complications (i.e., angina pectoris, heart failure, myocardial infarction, and stroke), complications of the lower extremities (neuropathy, including amputation and diabetic foot ulcer), eye problems (vision loss/blindness, cataract, and macular edema), end-stage renal disease, as well as hypoglycemia were included in most studies (>60%). Furthermore, a few studies included one or more of the following complications: depression, diabetic ketoacidosis, macro albuminuria, or coronary or ischemic heart disease. The latter complications overlap with other macrovascular complications such as myocardial infarction. Some CORE studies (7/24) also reported the inclusion of lactic acidosis, although the CORE Diabetes Model only applies this complication to the T2D population.

3.2 Baseline Characteristics and Continuous Glucose Monitoring Effects

Populations in the CUAs varied, ranging from all adults with T1D to populations with an increased risk for hypoglycemia, or individuals who did not achieve treatment goals for blood glucose levels. Only three studies included children and/or adolescents [39, 53, 61] (Table 1). Baseline characteristics varied widely across the studies. The baseline age and HbA1c values ranged from 12.0 to 51.4 years and 6.4 to 10.3% (46–89 mmol/mol), respectively (Online Resource 10). Several trials and observational studies were used to source the input data for the baseline characteristics and treatment effects (Table 2). The CGM effect was included in each study as a reduction in HbA1c (or time in

Fig. 2 Summary of the cost-effectiveness results of the included studies comparing to self-monitoring of blood glucose levels using finger pricks or other CGM devices. Note: Some of the studies reported multiple ICERs. The results for three studies [7, 28, 30] are not included because of missing incremental costs and/or QALY data. *CGM* continuous glucose monitoring, *ICERs* incremental cost-effectiveness ratios, *QALYs* quality-adjusted life-years, *SMBG* self-monitoring of blood glucose, *WTP* willingness to pay
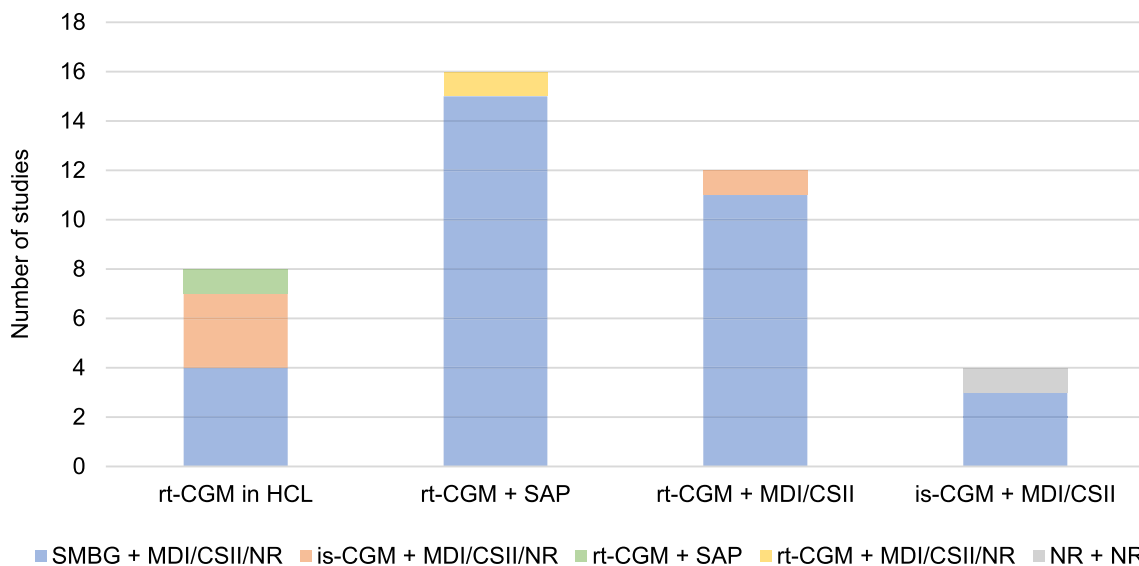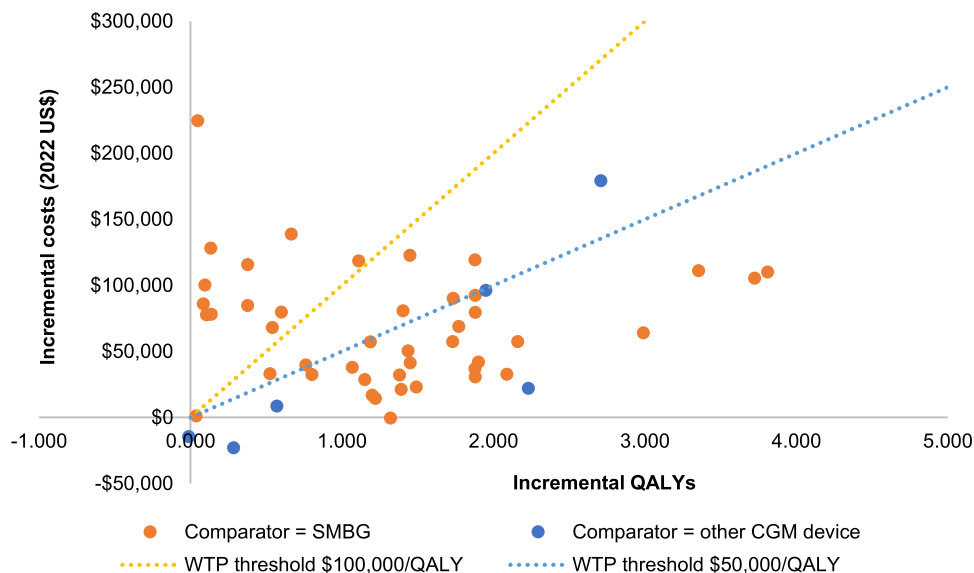

Fig. 3 Overview of the comparative analyses covered in the studies. Some studies included multiple comparisons. *CSII* continuous subcutaneous insulin infusion, *HCL* hybrid closed loop, *is-CGM* intermittently-scanned continuous glucose monitoring, *MDI* multiple daily injections, *NR* not reported, *rt-CGM* real-time continuous glucose monitoring, *SAP* sensor-augmented pump, *SMBG* self-monitoring of blood glucose

range [30]), and often also as a reduction in hypoglycemic events or in the fear of hypoglycemia, a reduction in ketoacidosis, or a combination of these. Finally, some studies modeled the direct effect on QoL associated with a reduced need for finger pricks [33].

Most studies used the same source for both the baseline characteristics and the CGM effect on HbA1c and hypoglycemic events. Commonly used sources were a meta-analysis by Pickup et al. [66] (7/34), and three RCTs: Ly et al. [67] (7/34), the DIAMOND trial [68, 69] (6/34), and

the FUTURE trial [70] (4/34). The remaining 10 studies used a variety of different RCTs. The effect on fear of hypoglycemia was calculated as a reduction in the score in the Hypoglycemia Fear Survey sourced from either the DIAMOND or the INTERPRET trial, or a meta-analysis by Yeh et al. [68, 69, 71, 72]. Reduction of ketoacidosis was only included in two studies [31, 50] and sourced from the FUTURE trial or a single-arm study (Table 2). Notably, the duration of the CGM effect was not reported in most of the studies (21/34), although we expect that these studies

assumed that CGM had a lifelong effect. Ten more studies explicitly reported using a (continuous) lifetime effect, without providing a justification. Only the health technology assessment (HTA) reports from the National Institute for Health and Care Excellence [36] and Health Quality Ontario [37] explicitly explained how the treatment effect was modeled over time. In these studies, a reduction in the treatment effect was assumed by using a continuous CGM effect in the first year, following which annual progression rates or a slowly declining effect were assumed. The authors justified this approach by the uncertainty of the effectiveness data, which showed considerable heterogeneity and differences in baseline HbA1c levels. Additionally, in both reports, a scenario analysis was conducted to estimate the effect of a lifetime treatment effect.

3.3 Costs

All the studies included costs related to diabetes treatments and complications. Those that applied a societal perspective all included productivity losses. However, only a few studies considered other societal costs such as costs related to training staff in CGM use and the time that patients/caregivers devoted to diabetes care [28, 54, 56]. Most studies used the human capital method to estimate costs related to losses in productivity.

The elements comprising intervention costs varied, and there was often lack of transparency about what was included and how this was valued. A number of studies only reported a lump-sum cost for the intervention and comparator, with some only reporting incremental costs without providing specific details. However, all 34 studies included the costs of CGM sensors (Online Resource 11). Costs of the transmitter/charger, test strips and lancets for SMBG (either as stand-alone comparators or for calibrating the CGM device) were only reported in 62% and 88% of the studies, respectively. Moreover, only 13/34 (38%) studies reported the costs of insulin pumps, and even fewer studies (10/34) considered other insulin pump-related costs, such as the costs of reservoir, infusion set, training, and pump replacement. The cost of insulin itself was included in 11/34 studies (32%). As costs could differ for the different devices and elements, the intervention costs were presumably often underestimated.

3.4 Health-Related Quality of Life Weights and Quality-Adjusted Life-Years

The studies that included analyses based on empirical data measured QoL in the trials themselves at baseline and at the end of each trial to obtain a CGM-related QALY benefit for use in the cost-effectiveness analysis [7, 28, 39, 56]. Two

of these studies also entailed a long-term analysis, adding modeling and QoL weights sourced from the literature. All except two [45, 49] of the model-based studies included QoL weights for macro- and microvascular complications. Fifteen studies on T1D populations used values derived from a review by Beaudet et al. [98], while other studies referred to a mix of studies. QoL effects for hypoglycemic events were included in 17/34 studies, with most of these studies (n = 9) citing the study by Evans et al. [99], who used time trade-off (TTO) in a general population with a mixed sample containing T1D and T2D populations from five countries. Incremental QoL gain was considered for the reduction in fear of hypoglycemia in 23/34 studies. All of these studies used a mapping of a 1-unit increase in the Hypoglycemia Fear Survey score to EQ-5D scores following the studies by Currie et al. and McBride et al. [73, 100] to convert Fear of Hypoglycemia scores into QALYs. Explicit QoL effects for ketoacidosis or reduction in finger pricks were rare, occurring in only 2/34 [30, 31] and 5/34 [38, 49, 59–61] studies, respectively.

Various methods were used to elicit QoL weights in the different referenced sources, although most applied EQ-5D (3L or 5L) or TTO. The populations for which complications were considered were the general population or individuals with T2D; those with T1D were seldom considered. The studies were conducted in diverse countries, and most models included values elicited in settings outside of their country contexts. While this approach is understandable given the large amount of QoL weights needed to reflect all diabetes-related complications included in the models, it deviates from the guidelines for economic evaluations. Online Resource 12 presents a summary of all sources and instruments used for QoL weights in the reviewed studies.

3.5 Quality Assessment

3.5.1 Quality of Model-Based Studies

Figure 4 and Online Resource 13 provide summaries of the findings using the Philips checklist. They show high scores for about half of the items (for 26/57 items, they were above 75%). These included items about choice of baseline data, synthesizing treatment effects (with many 'non-applicable' scores, as most studies applied a single source for their treatment effects), incorporation of appropriate utilities, addressing methodological uncertainty, and methods for assessing parameter uncertainty and reporting ranges for uncertainty. Scores for 19 items were below 20%. These items were: (1) 'Is the primary decision maker specified?'; (2) 'Have any competing theories regarding model structure been considered?'; (3) 'Is there justification for the exclusion of feasible options?'; (4) 'Is the duration of treatment and treatment

Table 2 Summary of sources used for baseline characteristics and continuous glucose monitoring effect size and duration

Publication (first author, year, country)	Source of baseline characteristics	Source of CGM effect size					Duration of CGM effect, motivation
		HbA1c reduction/time in range	Reduction in hypoglycemic events	Reduction in hypoglycemia	Reduction in fear of hypoglycemia	Reduction in ketoacidosis	
Emamipour (2022), The Netherlands [7]	Observational study (FLARE-NL4 study [5])	Observational study (FLARE-NL4 study [72])			–	–	12 months, same as study duration
Ly (2014), Australia [35]	RCT (Ly, 2013 [67])	–	RCT (Ly 2013 [67])		–	-	6 months, same as study duration
Wan (2018), USA [53]	RCT (DIAMOND trial [68])	Wan (2018), USA [53]			–	–	Within the trial analysis: 6 months; model-based analysis: lifetime, based on assumption
Bilir (2018), Sweden [45]	RCT (IMPACT trial [72])	–	RCT (IMPACT trial [72])		–	-	NR
Chaugule (2017), Canada [55]	RCT (DIAMOND trial [68])	RCT (DIAMOND trial [68])			Questionnaire (Currie, 2006 [73])	(Only in scenario analysis)	Lifetime
Conget (2018), Spain [30]	Unclear	–	Unclear		Observational study (INTERPRET study [72])	–	NR
Gómez (2016), Colombia [44]	Observational study (Gómez, 2013 [74])	Observational study (Gómez, 2013 [74])			Meta-analysis (Yeh, 2012 [75])	–	NR
Isitt (2022), Australia [34]	RCT (DIAMOND trial, [68])	RCT (DIAMOND trial [68]) + NMA (Pease, 2020 [76]) for is-CGM arm			RCT (DIAMOND trial [68])	–	NR
Jendle (2017), Sweden [35]	(Cohort 1) RCT (Ly, 2013 [67]), (Cohort 2) register (Swedish National Diabetes Register [77], supplemented by DCCT [78])	(Cohort 1) –; (Cohort 2) RCT (Ly, 2013 [67])	(Cohort 1) RCT (Ly, 2013 [67]); (Cohort 2) –		(Cohort 1) RCT (Ly, 2013 [67]) [75]; (Cohort 2) observational study (INTERPRET study [72])	–	NR
Jendle (2019), Sweden [51]	Single-arm, non-randomized trial (Bergenstal, 2016 [79])	Single-arm, non-randomized trial (Bergenstal, 2016 [79])			Observational study (INTERPRET study [72])	–	Lifetime
Jendle (2021), Sweden [54]	RCT (FUTURE study [70])	RCT (Collyns, 2021 [80])	RCT (FUTURE study [70])		Observational study (INTERPRET study [72])	–	Lifetime
Kamble (2012), USA [47]	RCT (STAR 3 study [81])	RCT (STAR 3 study [81])	–		–	–	Lifetime
Lambadiari (2022), Greece [33]	HCL vs. SAP plus PLGM: RCT (MiniMed 780G US pivotal trial [82]); AHCL vs. MDI plus is-CGM: observational study (FUTURE study [70])	RCT (Carlson, 2020 [83])					NR

Table 2 (continued)

Publication (first author, year, country)	Source of baseline characteristics	Source of CGM effect size				Duration of CGM effect, motivation
		HbA1c reduction/time in range	Reduction in hypoglycemic events	Reduction in fear of hypoglycemia	Reduction in ketoacidosis	
Nicolucci (2018), Italy [46]	(Cohort 1) RCT (Ly, 2013 [67]); (Cohort 2) meta-analysis (Pickup, 2011 [66])	(Cohort 1) –; (Cohort 2) meta-analysis (Pickup, 2011 [66])	(Cohort 1) RCT (Ly, 2013 [67]); (Cohort 2) –	(Cohort 1) RCT (Ly, 2013, [67]);(Cohort 2) observational study (INTERPRET study [72])	–	NR
Riemsma (2016), UK [36]	Single-arm, non-randomized trial (Bergenstal, 2016 [79]) supplemented by various sources	Self-conducted meta-analysis		(Only in sensitivity analysis)	–	Only the first year; an annual progression rate is subsequently applied
Roze (2015), Sweden [40]	Meta-analysis (Pickup, 2011 [66])	Meta-analysis (Pickup, 2011 [66])	(Only in sensitivity analysis)	RCT (DIAMOND trial, [68])	–	NR
Roze (2016), France [41]	(Cohort 1) RCT (Riveline, 2012 [84]); (Cohort 2) RCT (Ly, 2013 [67])	(Cohort 1) Meta-analysis (unpublished); (Cohort 2) –	(Cohort 1) –; (Cohort 2) RCT (Ly, 2013 [67])	Observational study (INTERPRET study [72])	–	NR
Roze (2016), UK [42]	Meta-analysis (Pickup, 2011 [66])	Meta-analysis (Pickup, 2011 [66])	Observational study (Lynch, 2012 [85])	Observational study (INTERPRET study [72])	–	NR
Roze (2017), Denmark [44]	(Cohort 1) Meta-analysis (Pickup, 2011 [66]); (Cohort 2) RCT (Ly, 2013, [67])	(Cohort 1) Meta-analysis (Pickup, 2011 [66]); (Cohort 2) –	(Cohort 1) –; (Cohort 2) RCT (Ly, 2013 [67])	(Cohort 1) Observational study (INTERPRET study [72]); (Cohort 2) RCT (Ly, 2013 [67])	–	NR
Roze (2019), the Netherlands [45]	(Cohort 1) Meta-analysis (Pickup, 2011 [66] supplemented by DCCT, [78]); (Cohort 2) RCT (Ly, 2013 [67])	(Cohort 1) Meta-analysis (Pickup, 2011 [66]); (Cohort 2) –	(Cohort 1) –; (Cohort 2) RCT (Ly, 2013 [67])	(Cohort 1) Observational study (INTERPRET study [72]); (Cohort 2) RCT (Ly, 2013 [67])	–	NR
Roze (2019), Turkey [43]	(Cohort 1) Meta-analysis (Pickup, 2011 [66]); (Cohort 2) RCT (Ly, 2013 [67])	(Cohort 1) Meta-analysis (Pickup, 2011 [66]); (Cohort 2) –	(Cohort 1) –; (Cohort 2) RCT (Ly, 2013 [67])	(Cohort 1) Observational study (INTERPRET study [72]); (Cohort 2) RCT (Ly, 2013 [67])	–	NR
Roze (2020), UK [59]	RCT (DIAMOND trial [68])	RCT (DIAMOND trial [68])	RCT (DIAMOND trial [68])	RCT (DIAMOND trial, [68])	–	NR
Roze (2021), Canada [60]	RCT (DIAMOND trial [68])	RCT (DIAMOND trial [68])	RCT (DIAMOND trial [68])	RCT (DIAMOND trial, [68])	–	Lifetime
Roze (2021), UK [50]	RCT (DIAMOND trial [68])	Single-arm, non-randomized trial (Garg, 2017/Bergenstal, 2016 [79, 86])	Single-arm, non-randomized trial (Garg, 2017/Bergenstal, 2016 [79, 86])	Observational study (INTERPRET study [72])	Single-arm, non-randomized trial (Garg, 2017/ Bergenstal, 2016 [79, 86])	NR

Table 2 (continued)

Publication (first author, year, country)	Source of baseline characteristics	Source of CGM effect size				Duration of CGM effect, motivation
		HbA1c reduction/time in range	Reduction in hypoglycemic events	Reduction in fear of hypoglycemia	Reduction in ketoacidosis	
Roze (2021), France [61]	Single-arm, non-randomized trial (Garg, 2017/Bergenstal, 2016 [79, 86])	RCT (DIAMOND trial [68])	–	RCT (DIAMOND trial [69])	–	Lifetime
Serné (2022), the Netherlands [55]	RCT (FUTURE study [70])	RCT (FUTURE study [70])		Observational study (INTERPRET study [72])	–	NR
Zhao (2021), China [31]	Epidemiological studies (Zhou, 2020 and Tang, 2019 [87, 88])	RCT (IMPACT trial [89])		–	RCT (FUTURE study [70])	NR
Garcia-Lorenzo (2018), Spain [57]	Meta-analysis targeting the study objectives	Meta-analysis (Pickup, 2011 [66])	–	–	–	NR
Health Quality Ontario (2018), Canada [37]	RCT (DCCT [78])	CGM + MDI vs. SMBG + MDI: RCT (GOLD trial [90]); SAP vs. SMBG + MDI: RCT (STAR3 trial, [81]); SAP vs. SMBG + CSII: RCT (Tumminia, 2015 [91])	All comparisons: RCT (STAR 3 study [81])	–	–	Only the first year, subsequently declining slowly; assumption-based
Huang (2010), USA [28]	RCT (Juvenile Diabetes Research Foundation-CGM trial [92])	RCT (Juvenile Diabetes Research Foundation-CGM trial [92])	–	–	–	Lifetime; assumption-based
McQueen (2011), USA [32]	RCT (Juvenile Diabetes Research Foundation-CGM trial [92])	RCT (Juvenile Diabetes Research Foundation-CGM trial [92])	–	–	–	NR
Pease (2020), Australia [52]	Register (Australian National Diabetes Audit [93])	RCT (Thabit, 2015 [94])	Severe: assumption-based; non-severe: RCT (Ben-hamou, 2019 [95])	–	–	NR
Pease (2022), Australia [53]	RCT (Abraham, 2021 [96])	RCT (Abraham, 2021 [96])	Severe: assumption-based + other CEA (Pease, 2020 [52]); non-severe: RCT (Abraham, 2020 [96]) + other CEA (Pease, 2020 [52])	–	–	Continuous over a 13-year time horizon; assumption-based
Rotondi (2022), Canada [30]	NR	Meta-analysis (Maiorino, 2020 [97])	–	–	–	NR

AHCL advanced hybrid closed loop, *CEA* cost-effectiveness analysis, *CGM* continuous glucose monitoring, *CSII* continuous subcutaneous insulin infusion, *DCCT* Diabetes Control and Complications Trial, *HCL* hybrid closed loop, *is-CGM* intermittently-scanned continuous glucose monitoring, *MDI* multiple daily injections, *NMA* network meta-analysis, *NR* not reported, *PLGM* predictive low glucose monitoring, *RCT* randomized controlled trial, *SAP* sensor-augmented pump, *SMBG* self-monitoring of blood glucose

Model quality: Philips, scores per item

Fig. 4 Summary of quality scores for the Philips checklist. Descriptions are short versions of the full-item texts, see Online Resource 3 for full-item descriptions and how these were operationalized by the authors. Note: Online Resource 11 provides an explanation of the items. *NA* not applicable, *NR* not reported, *Y/N* 'yes' or 'no' unclear
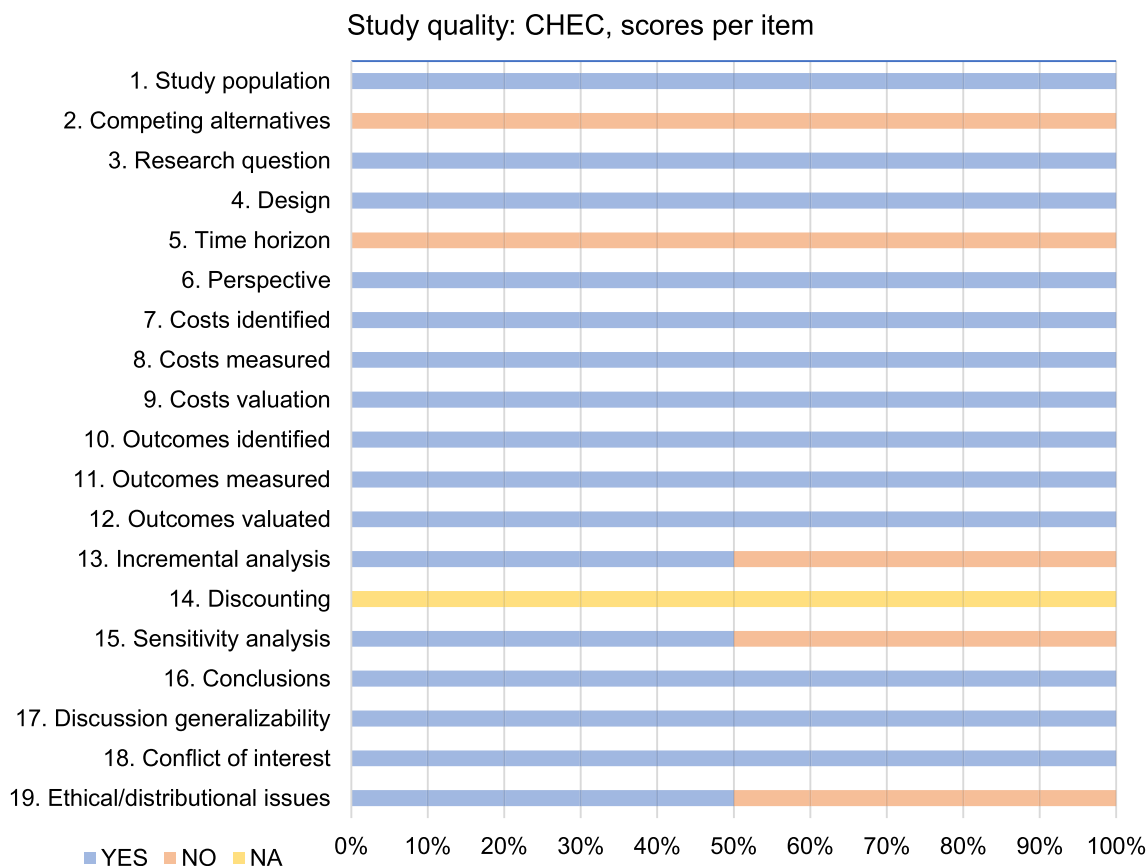
Study quality: CHEC, scores per item

Fig. 5 Model quality in studies using empirical data and the CHEC-extended checklist [22, 23]. *CHEC* Consensus Health Economic Criteria, *NA* not applicable

effect described and justified?'; (5) 'Is there evidence that the mathematical logic of the model has been tested thoroughly before use?'; and another 14 that concerned data identification and analysis, modeling details on baseline and treatment effects, data incorporation approaches, and uncertainty analysis. In general, the studies lacked a description of systematic and justified choices in the modeling process. Furthermore, most studies only included univariate sensitivity analyses, while probabilistic sensitivity analyses were not always performed. The Philips quality scores per paper were moderate (on average, 56%), ranging between 39% and 87% (Online Resource 13).

3.5.2 Quality of Studies Using Empirical Data

Figure 5 and Online Resource 14 present scores for the two studies that only used empirical data [7, 39]. The discrepancy scores for disagreements between authors regarding extracted data were much lower for the CHEC-extended list (21% of the items) than those for the Philips checklist (74% initially). The overall score was high (a 'yes' response for 76% of the items), with comparable scores for

the two studies. Both studies scored a 'no' for the question on whether all alternatives were clearly described. The time horizons in both studies were insufficient to include relevant costs and consequences, as the studies did not consider the long-term complications of CGM use beyond the trial period. Given the short time horizon, item 14 ('Are all future costs and outcomes discounted appropriately?') was not applicable.

3.6 Model Validation

Figure 6 shows the scores for the model validation items in AdViSHE (see Online Resource 15 for details provided in each paper). In general, the authors paid scant attention to model validity, with most simply stating that a validated model was used without providing further details. However, all but one study that applied the CORE Diabetes Model referred to the original studies reporting on the structure of the model and its validation [62–64]. Hence, we assigned them similar scores for model validity based on these references. This approach impacted the scores provided below. Therefore, we also reported scores

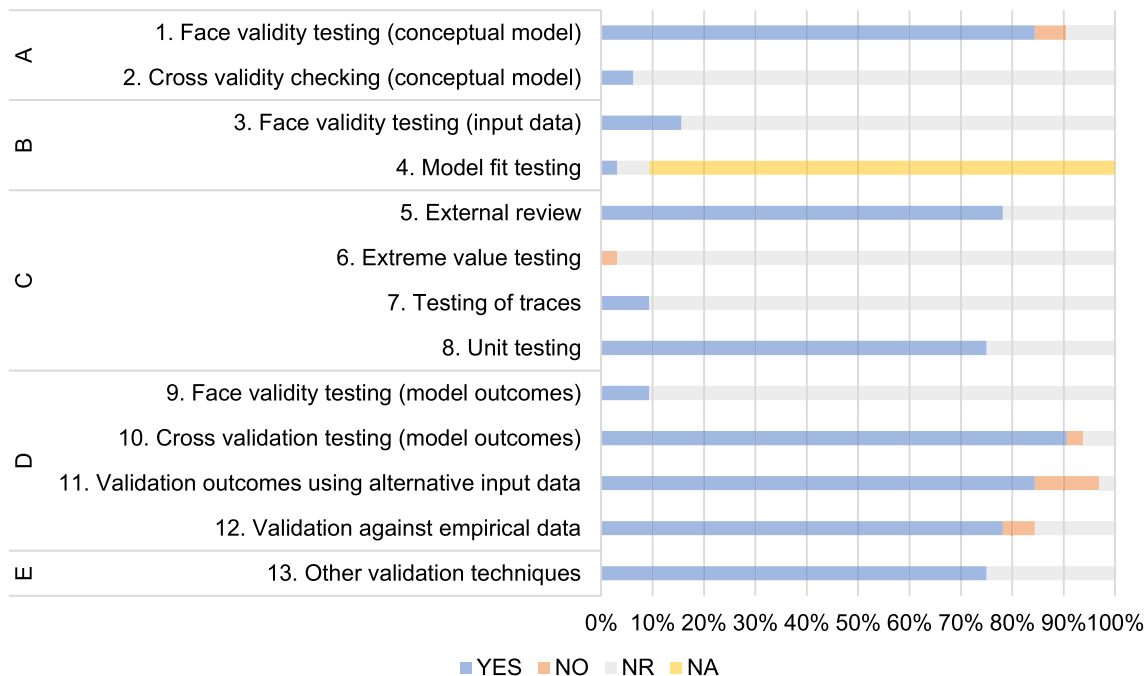## Model validation: AdViSHE, scores per item

Fig. 6 Scores for model validation tests performed and reported in the papers, supplemental information, or direct references for the 32 model-based economic evaluations. Notes: Papers applying the CORE Diabetes Model that referred to the study by McEwan et al. 2014 [64], scored 'yes' for cross-validation tests (2 and 10), comparing model outcomes with empirical data (12), testing with alternative input data (11), and other tests (13). This is because the model has undergone validation within the Mount Hood Diabetes Challenge Network. The same situation applies to the Sheffield model [56]. *AdViSHE* Assessment of the Validation Status of Health-Economic decision models, *NA* not applicable, *NR* not reported

on a model-level (Online Resource 15). Items that were reported most often focused on face validity checking of the conceptual model (26/32, 81%); expert checks of the simulation code (25/32, 78%); comparisons of the model outcomes with those of other models (25/32, 78%), and the use of alternative sets of input data (27/32, 84%). Comparisons made with other models were often basic, comprising a short, qualitative comparison in the Discussion section. For some of the included models, we know from conferences and publications that cross-model validation was performed, including comparisons with empirical data and searches for alternative sets of input data [101–103]; however, these validation procedures were not mentioned in the publications. Items with consistently low scores across all studies were cross-model validation of the conceptual model and code verification by extreme value testing. It is possible that such tests, although not reported in most studies, were performed. An additional validation effort reported was the recoding of several elements of the model in Microsoft Excel, indicating the complexity of diabetes models and the need to check model codes carefully [63, 64].

4 Discussion

4.1 Main Study Results

CUAs of CGM showed mixed results for quality and transparency. Considerable improvements, especially for increasing transparency, are possible. The supplementary tables included in this study presenting detailed information on all the studies may aid decision makers in various settings to determine which studies are relevant to them, so that they do not have to examine each study individually. With few exceptions, most studies would be very hard, if not impossible, to replicate, even considering the content of references and online supplementary information.

This finding confirms that of a previous attempt to replicate the results of two published studies using the input data described in those articles by modeling groups coordinated by the Mount Hood Diabetes Challenge Network [101]. The Mount Hood group used this study to formulate specific guidelines to promote transparency in cost-effectiveness studies focusing on diabetes [101]. Most

of the studies in our review would not meet these standards. In particular, the costing of interventions, how effects were modeled, and assumptions on HbA1c trajectories were poorly reported. Another interesting advice in this guideline with potential cost-effectiveness implications concerns the choice of mortality table and event-specific mortality. Relevant items to address this were not explicitly stated in our data extraction tables and merit attention in future studies. Authors, reviewers, and editors alike may need to attend to the issue of transparency and could benefit from the use of disease-specific as well as general cost-effectiveness reporting guidelines and checklists to enhance transparency and replicability.

Other recent reviews that critically appraise health economic models have pointed to important omissions in model quality and validation similar to those that we identified. For instance, recent studies on cancer screening and mental illness indicate a lack of validation and express model quality concerns [104, 105]. Structural uncertainty analyses and explanation of choices regarding input data were explicitly noted as missing in studies on mental illness. Given the option of online supplementary sections, the argument of a lack of space is untenable and cannot justify omitting important information on the model version and parameters applied or about assumptions regarding long-term effects. Therefore, we recommend that authors of model-based economic evaluation studies provide more detailed reporting and use quality checklists to verify that important elements have been addressed.

Our review complements a recent review on CGM evaluations by Jiao et al. [8]. Our findings, like those of that study, showed diverse outcomes with significant differences among ICERs, even for studies that compared the same or similar devices. However, we examined the quality of studies, input data sources, and model validity in more depth. We also covered 15 studies not included in the earlier review. Of these, nine were published after the timeline covered by that review, which confirms that CGM in T1D is a rapidly changing field.

Given the chronic nature of T1D and the possible preventive effects of CGM on diabetes complications in the long run, evidence of the cost utility of CGM in T1D primarily relies on short-term clinical trial results, which are extrapolated over an individual's lifetime using health economic models. We only identified two stand-alone studies based on empirical data; two more studies reported short-term within-trial results in combination with lifetime model-based outcomes. The two stand-alone studies obtained good CHEC scores but only included health benefits over a limited period (6–12 months). Thus, the potential long-term effects of improved glycemic control on diabetes-related complications were not explored. Nevertheless, both studies found

CGM to be cost effective. By contrast, ICERs ranged from dominant to values over US$4 million per QALY in the model-based studies. The magnitude of short-term benefits found in these two studies were 0.030–0.037 QALYs. These values lie within the − 0.014 to 3.810 range reported for the model-based studies.

4.2 Critical Appraisal

The majority of the studies were based on models. In this section, we critically appraise key findings on the choice of comparator, model quality and validation, the modeling of effects and costs, and QoL to provide a rationale for our recommendations.

The reviewed studies evaluated a variety of combinations of modes of insulin administration and CGM devices. To remain relevant in a field characterized by rapid innovation, the choice of relevant comparators is important. These should comprise the next best alternative, or care-as-usual options. However, many studies compared a single combination of a CGM device and insulin administration with SMBG. In many high-income countries, SMBG may no longer be a relevant comparator, with this comparison yielding no insight into second-best alternatives. Given the wide variety of available options, a systematic approach is required. This was only accomplished in two extensive HTA reports [36, 37] commissioned by a healthcare decision maker. Ideally, studies should entail a (network) meta-analysis of all relevant alternatives to assess the effectiveness of various treatment options [21].

Although CGM entails lifelong blood glucose management, effects have only been measured in studies with limited follow-up times [71, 106]. Studies with a follow-up period of 2 years showed that patients experienced a sustained improvement in glycemic control and QoL [4]. The conclusions of systematic reviews regarding effects other than HbA1c reduction were conservative [71, 106]. While all studies used HbA1c reduction as the primary CGM effect, most studies also included fear of hypoglycemia as an effect. Only two studies included an effect on ketoacidosis [31, 50]. The studies in our review generally did not report the duration of CGM effects; those that did so using a lifetime (continuous) effect offered no justification. Only two studies considered the possibility that effects would change over time [36, 37]. In sum, most studies were unclear or overly optimistic about long-term effects on the one hand, while on the other hand, a few studies used conservative assumptions that in turn could be overcautious. Future evaluations of CGM should refer to available meta-analyses and make transparent and reasonable assumptions regarding the duration of the CGM effect. The effects of alternative assumptions concerning the duration of the effect and impact

on hypoglycemia, fear of hypoglycemia, and ketoacidosis should be explored using elaborate sensitivity analyses.

The QoL used in the studies often included a mix of sources, including T2D and general population studies using various questionnaires administered in different countries. For complications other than hypoglycemia, fear of hypoglycemia, and ketoacidosis, it is probably acceptable to use QoL weights from non-T1D populations, but some justification would be expected. For uncomplicated T1D (baseline utility) and complications directly affected by glucose monitoring, the use of specific QoL studies linked to T1D is probably optimal.

Cost-analysis approaches varied widely, and costing elements evidenced little consistency. This is a concern given the existence of clear guidelines [19]. In our opinion, costing should be complete, and should always include the costs of monitoring devices, insulin administration devices and consumables such as test strips as well as eventual device replacement. Moreover, insulin costs were only included for one-third of the studies, although insulin use and therefore related costs can be affected by CGM.

In light of these findings, the considerable variability in cost-effectiveness outcomes reported in a previous review [8] can be partly explained by substantial differences in costing approaches as well as differences in the types of interventions, populations, and corresponding effect sizes. However, differences in the models used likely explain little of these differences, as most of the selected studies applied the CORE Diabetes Model or a variant of the model developed by McQueen et al. [32].

Important omissions apart from the above-discussed lack of clarity concerning extrapolation of treatment effects included the lack of justification of model and input data choices and of a sound structural uncertainty analysis. These all reduced study quality scores. Although several studies analyzed parameter uncertainty, quality and transparency were often problematic. For instance, only a selection of parameters was varied in the univariate sensitivity analysis, and the information on distributions used was often incomplete. Moreover, while most studies included some form of deterministic sensitivity analysis, not all studies included a state-of-the-art probabilistic sensitivity analysis [22]. All of these gaps underline the importance of following existing model quality and reporting guidelines provided in the Philips checklist.

In diabetes modeling, an active group of modelers perform cross-model comparisons and address challenges relating to internal and external validation [101–103]. Therefore, we expected diabetes models to stand out in this area. However, limited reporting of validation procedures was performed in the reviewed papers, and the few available references were limited to validation using older-model versions. Furthermore, validation of conceptual models or input data was poorly reported and likely underperformed in most studies included in our review. This issue was linked to low scores for justification of the model structure or structural uncertainty testing using the Philips checklist. It was also linked to many items on input data choices and processing. Validation of both the conceptual model and input data and good modeling practice should therefore receive more attention in future studies.

Referencing a dedicated paper on model validation is not sufficient in itself, as a model should be valid for its intended application. Accordingly, it should be sufficiently adapted to the application setting [29]. Few studies tested whether the intervention effect sizes were validly implemented into the model. Moreover, few studies reported on any setting-specific adaptation other than the application of local costs. Additionally, applications should use the most recent model version. However, of the 24 studies using the CORE Diabetes Model, few reported the model version, and the studies invariably referred to early (2004) CORE publications, which do not elucidate current T1D applications.

4.3 Limitations and Strengths of the Study

Limitations of this review relate to the inherently subjective nature of quality and transparency scoring. In particular, the Philips checklist raised many issues, with an initial high degree of disagreement among the four co-authors who scored these items. Even after several consensus-building meetings, inconsistencies remained between scoring duos. We solved this issue by rescoring after including additional agreements about scoring an item as being of sufficient quality. We attempted to strike a balance between stringency and leniency. Being overly strict and only judging what was reported in the paper itself, or its supplementary sections, would have resulted in negative responses for most items in almost all of the studies. Therefore, we included references to key modeling studies, and for a few items we scored quality at the level of the model rather than the paper. Similarly, our AdViSHE scores were model-centered. Had we based them purely on the evaluation studies, the scores would have been very low. Having 24 studies of a single model, compared with eight studies with other models, will have impacted our results for model quality and validity. For model validity, we added a figure with scores at model level to the online resources, indicating this would not substantially change conclusions. For Philips, this was not a feasible option, given that many items refer to the application, not to the model.

Although we used an approach of independent scoring by two authors, followed by consensus-building meetings, and checked both references and supplemental information,

scoring quality on the basis of publications has limitations. This is because studies could have performed well for some aspects, which the authors did not report. In such cases, we could not appreciate the quality of the complete study. However, as we used the well-known Philips and AdViSHE checklists, it could be argued that lack of clarity on items in these checklists also indicated poor quality, if not in the modeling or validation approach then in the reporting.

It should be noted that decision models in health economics require maintenance and updates. Therefore, it is questionable whether some model-based studies were referring to the same model version as the one applied in the economic evaluation, which will have affected our scores. Another limitation was that data extraction was performed by four co-authors, implying that no single author extracted all the studies. However, the advantage of our approach was that it highlighted inconsistencies between duos, and we held several consensus-building meetings, thereby reducing the risk of bias in scoring subjective items.

A strength of this review was that we paid special attention to all costing elements and QoL weights included in the various studies. Furthermore, we examined methodological quality and model validation. In addition, we carefully sorted out the diverse devices and classified them by monitoring device and insulin administration methods. A co-author with clinical expertise checked all classifications.

The comparability of study outcomes could potentially be reduced by the broad range of selection criteria that we applied. However, the goal of our review was not to explain differences in ICERs between specific studies; rather, it was to analyze the variety of methodological decisions, which could affect not only results but also the quality of the analyses. Strict selection criteria relating to the study population, insulin administration, ethnicities, nationalities, and age could result in less diverse ICERs. However, such a selection would not be a random sample in terms of methodological quality and give an incomplete picture of quality.

4.4 Summary of Recommendations

We have presented a critical appraisal of the studies' performance relating to transparency, study quality, and model validation. We did not identify any studies that could serve as models of good practice for all these elements, although the two HTA reports performed well in many respects [36, 37]. We observed a large set of routinely performed repetitions of similar comparisons in many settings among the included studies. Interpreting the diverse outcomes and relating them to differences in model quality was challenging given the lack of transparency. However, we tagged certain points, consideration of which could lead to favorable ICERs. The first was the assumption of an effect on fear of hypoglycemia despite unclear evidence for this in recent

meta-analyses [71, 106]. The second concerns the exclusion of certain cost components, such as changes in insulin use and insulin pump-related costs. The third is the assumption of lifelong effects on glucose levels in the absence of trials with a follow-up period beyond 2 years. Last, most interventions were compared with SMBG, even for sophisticated CGM devices that could also be compared with simpler devices instead of SMBG.

The main recommendation for researchers performing economic evaluations of CGM is that they should follow existing guidelines on best practices concerning model transparency and validation, and use the available reporting tools. Several reporting checklists are relevant, including CHEERS [19], the Diabetes transparency checklist [96], the Philips checklist [21] and the AdViSHE checklist [24].

An 'ideal' study could use the following specific recommendations. First, studies should apply a (network) meta-analysis of all relevant alternatives to assess the effectiveness of various treatment options [21]. Second, assumptions on CGM effects and their duration should be transparent and reasonable, and at the very least, elaborate sensitivity analyses should be performed using available meta-analyses [71, 106]. Third, we recommend the use of specific QoL studies for assessments of the uncomplicated stage of T1D and for complications directly affected by glucose monitoring. Fourth, costing should be complete. Fifth, researchers should pay more attention to the quality of the conceptual model and input data, justifying their choices. Last, models should be sufficiently validated in terms of the intervention at stake to ensure their adaptation to the setting. Moreover, the model version should be reported.

5 Conclusion

CUAs of CGM showed deficits in quality and transparency. Improvements are possible through the adoption of systematic approaches to data selection, performance of uncertainty analyses, enhanced clarity of reporting, and the justification of choices made in the modeling process. The requirement of comparing all relevant and feasible options was rarely met, while this is essential to assure decision makers, especially in the case of a rapidly evolving technology such as CGM and insulin administration.

Most of the decision models used in health economics have been well documented in background papers, and some have been extensively validated. Our review revealed that this is a necessary condition for replicable and good-quality applications but it is not sufficient in itself, highlighting the need to follow existing guidelines on best practices.

Declarations

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication (from patients/participants) Not applicable.

Code Availability Not applicable.

References

1. Nathan D, Genuth S, Lachin J, Cleary P, Crofford O, Davis M, et al. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med. 1993;329:977–86.

2. Turner R. Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). Lancet. 1998;352:854–65.

3. Song MY, Steinhubl SR, Topol EJ. The beginning of the end of fingersticks? The Lancet. 2018;392:203.

4. Lameijer A, Fokkert MJ, Edens MA, Gans ROB, Bilo HJG, Van Dijk PR. Two-year use of flash glucose monitoring is associated with sustained improvement of glycemic control and quality of life (FLARE-NL-6). BMJ Open Diabetes Res Care. 2021;9(1):e002124.

5. Fokkert M, Van Dijk P, Edens M, Barents E, Mollema J, Slingerland R, et al. Improved well-being and decreased disease burden after 1-year use of flash glucose monitoring (FLARE-NL4). BMJ Open Diabetes Res Care. 2019;7(1):e000809.

6. Holt RIG, DeVries JH, Hess-Fischl A, Hirsch IB, Kirkman MS, Klupa T, et al. The management of type 1 diabetes in adults. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetologia. 2021;64:2609–52.

7. Emamipour S, van Dijk PR, Bilo HJG, Edens MA, van der Galiën O, Postma MJ, et al. Personalizing the Use of a Intermittently Scanned Continuous Glucose Monitoring (isCGM) Device in Individuals With Type 1 Diabetes: A Cost-Effectiveness Perspective in the Netherlands (FLARE-NL 9). J Diabetes Sci Technol. 2024;18(1):135–42.

8. Jiao Y, Lin R, Hua X, Churilov L, Gaca MJ, James S, et al. A systematic review: COST-effectiveness of continuous glucose monitoring compared to self-monitoring of blood glucose in type 1 diabetes. Endocrinol Diabetes Metab. 2022;5(6): e369.

9. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015;2015:350.

10. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4:148–60.

11. PROSPERO. Available at: https://www.crd.york.ac.uk/prospero/#searchadvanced. Accessed 18 Oct 2022.

12. de Jong LA, Li X, Emamipour S, van der Werf S, Postma MJ, van Dijk PR, et al. Model and empirical data-based cost-utility studies of continuous glucose monitoring in individuals with type 1 diabetes: a protocol of a systematic review on methodology and quality. Pharmacoecon Open. 2023;7(6):1007–13.

13. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, The PRISMA, et al. statement: an updated guideline for reporting systematic reviews. BMJ. 2020;2021:372.

14. Reddy N, Verma N, Dungan K. Monitoring Technologies- Continuous Glucose Monitoring, Mobile Technology, Biomarkers of Glycemic Control. Endotext. 2020; South Dartmouth (MA): MDText.com, Inc.; 2000.

15. Girardin CM, Huot C, Gonthier M, Delvin E. Continuous glucose monitoring: a review of biochemical perspectives and clinical use in type 1 diabetes. Clin Biochem. 2009;42:136–42.

16. Dicembrini I, Cosentino C, Monami M, Mannucci E, Pala L. Effects of real-time continuous glucose monitoring in type 1 diabetes: a meta-analysis of randomized controlled trials. Acta Diabetol. 2020;58:401–10.

17. Bramer WM, Giustini D, De Jong GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. J Med Libr Assoc. 2016;104:240–3.

18. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev. 2016;5:210.

19. Husereau D, Drummond M, Augustovski F, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Explanation and Elaboration: A report of the ISPOR CHEERS II Good Practices Task Force. Value in Health. 2022;25:10–31.

20. Wijnen B, Van Mastrigt G, Redekop W, Majoie H, De Kinderen R, Evers SMAA. How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: data extraction, risk of bias, and transferability (part 3/3). Expert Rev Pharmacoecon Outcomes Res. 2016;16:723–32.

21. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics. 2006;24:355–71.

22. Odnoletkova I. Cost-effectiveness of therapeutic education to prevent the development and progression of type 2 diabetes: systematic review. J Diabetes Metab. 2014;5:9

23. Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic

evaluations: Consensus on Health Economic Criteria - Pub-Med. Int J Technol Assess Health Care. 2005;21:240–5.

24. Vemer P, Corro Ramos I, van Voorn GAK, Al MJ, Feenstra TL. AdViSHE: a validation-assessment tool of health-economic models for decision makers and model users. Pharmacoeconomics. 2016;34:349.

25. OECD. Purchasing power parities (PPP). 2024. Available at: https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm#indicator-chart. Accessed 29 Jan 2024.

26. OECD. Exchange rates. 2024. Available at: https://data.oecd.org/conversion/exchange-rates.htm#indicator-chart. Accessed 29 Jan 2024.

27. Turner HC, Lauer JA, Tran BX, Teerawattananon Y, Jit M. Adjusting for inflation and currency changes within health economic studies. Value Health. 2019;22:1026–32.

28. Huang ES, O'Grady M, Basu A, Winn A, John P, Lee J, et al. The cost-effectiveness of continuous glucose monitoring in type 1 diabetes. Diabetes Care. 2010;33:1269–74.

29. Emamipour S, Pagano E, Di Cuonzo D, Konings SRA, van der Heijden AA, Elders P, et al. The transferability and validity of a population-level simulation model for the economic evaluation of interventions in diabetes: the MICADO model. Acta Diabetol. 2022;59:949–57.

30. Rotondi MA, Wong O, Riddell M, Perkins B. population-level impact and cost-effectiveness of continuous glucose monitoring and intermittently scanned continuous glucose monitoring technologies for adults with type 1 diabetes in Canada: a modeling study. Diabetes Care. 2022;45:2012–9.

31. Zhao X, Ming J, Qu S, Li HJ, Wu J, Ji L, et al. Cost-effectiveness of flash glucose monitoring for the management of patients with type 1 and patients with type 2 diabetes in China. Diabetes Ther. 2021;12:3079–92.

32. McQueen RB, Ellis SL, Campbell JD, Nair K V., Sullivan PW. Cost-effectiveness of continuous glucose monitoring and intensive insulin therapy for type 1 diabetes. Cost Eff Resour Alloc. 2011;9:13.

33. Lambadiari V, Ozdemir Saltik AZ, De Portu S, Buompensiere MI, Kountouri A, Korakas E, et al. Cost-effectiveness analysis of an advanced hybrid closed-loop insulin delivery system in people with type 1 diabetes in Greece. Diabetes Technol Ther. 2022;24:316–23.

34. Conget I, Martín-Vaquero P, Roze S, Elías I, Pineda C, Álvarez M, et al. Cost-effectiveness analysis of sensor-augmented pump therapy with low glucose-suspend in patients with type 1 diabetes mellitus and high risk of hypoglycemia in Spain. Endocrinol Diabetes Nutr. 2018;65:380–6.

35. Jendle J, Smith-Palmer J, Delbaere A, de Portu S, Papo N, Valentine W, et al. Cost-effectiveness analysis of sensor-augmented insulin pump therapy with automated insulin suspension versus standard insulin pump therapy in patients with type 1 diabetes in Sweden. Diabetes Ther. 2017;8:1015–30.

36. Riemsma R, Ramos IC, Birnie R, Büyükkaramikli N, Armstrong N, Ryder S, et al. Integrated sensor-augmented pump therapy systems [the MiniMed® Paradigm™ Veo system and the Vibe™ and G4® PLATINUM CGM (continuous glucose monitoring) system] for managing blood glucose levels in type 1 diabetes: a systematic review and economic evaluation. Health Technol Assess. 2016;20:1–252.

37. Quality Health Ontario. Continuous Monitoring of Glucose for Type 1 Diabetes: A Health Technology Assessment. Ontario; 2018.

38. Isitt JJ, Roze S. Long-term cost-effectiveness of Dexcom G6 real-time continuous glucose monitoring system in people with type 1 diabetes in Australia: response to letter from Hellmund, Richard and Welsh, Zoe. Diabet Med. 2022;39(7):e14831.

39. Ly TT, Brnabic AJM, Eggleston A, Kolivos A, McBride ME, Schrover R, et al. A cost-effectiveness analysis of sensor-augmented insulin pump therapy and automated insulin suspension versus standard pump therapy for hypoglycemic unaware patients with type 1 diabetes. Value Health. 2014;17:561–9.

40. Roze S, Saunders R, Brandt AS, de Portu S, Papo NL, Jendle J. Health-economic analysis of real-time continuous glucose monitoring in people with Type 1 diabetes. Diabet Med. 2015;32:618–26.

41. Roze S, Smith-Palmer J, Valentine W, Payet V, De Portu S, Papo N, et al. Cost-effectiveness of sensor-augmented pump therapy with low glucose suspend versus standard insulin pump therapy in two different patient populations with type 1 diabetes in France. Diabetes Technol Ther. 2016;18:75–84.

42. Roze S, Smith-Palmer J, Valentine WJ, Cook M, Jethwa M, De Portu S, et al. Long-term health economic benefits of sensor-augmented pump therapy vs continuous subcutaneous insulin infusion alone in type 1 diabetes: a U.K. perspective. J Med Econ. 2016;19:236–42.

43. Roze S, Smith-Palmer J, De Portu S, Özdemir Saltik AZ, Akgül T, Deyneli O. Cost-effectiveness of sensor-augmented insulin pump therapy versus continuous insulin infusion in patients with type 1 diabetes in Turkey. Diabetes Technol Ther. 2019;21:727–35.

44. Roze S, de Portu S, Smith-Palmer J, Delbaere A, Valentine W, Ridderstråle M. Cost-effectiveness of sensor-augmented pump therapy versus standard insulin pump therapy in patients with type 1 diabetes in Denmark. Diabetes Res Clin Pract. 2017;128:6–14.

45. Roze S, Smith-Palmer J, De Portu S, Delbaere A, De Brouwer B, De Valk HW. Cost-effectiveness of sensor-augmented insulin pump therapy vs continuous subcutaneous insulin infusion in patients with type 1 diabetes in the Netherlands. Clinicoecon Outcomes Res. 2019;11:73–82.

46. Nicolucci A, Rossi MC, D'Ostilio D, Delbaere A, de Portu S, Roze S. Cost-effectiveness of sensor-augmented pump therapy in two different patient populations with type 1 diabetes in Italy. Nutr Metab Cardiovasc Dis. 2018;28:707–15.

47. Kamble S, Schulman KA, Reed SD. Cost-effectiveness of sensor-augmented pump therapy in adults with type 1 diabetes in the United States. Value Health. 2012;15:632–8.

48. Gomez AM, Alfonso-Cristancho R, Orozco JJ, Lynch PM, Prieto D, Saunders R, et al. Clinical and economic benefits of integrated pump/CGM technology therapy in patients with type 1 diabetes in Colombia. Endocrinol Nutr. 2016;63:466–74.

49. Bilir SP, Hellmund R, Wehler B, Li H, Munakata J, Lamotte M. Cost-effectiveness analysis of a flash glucose monitoring system for patients with type 1 diabetes receiving intensive insulin treatment in Sweden. Eur Endocrinol. 2018;14:73–9.

50. Roze S, Buompensiere MI, Ozdemir Z, de Portu S, Cohen O. Cost-effectiveness of a novel hybrid closed-loop system compared with continuous subcutaneous insulin infusion in people with type 1 diabetes in the UK. J Med Econ. 2021;24:883–90.

51. Jendle J, Pöhlmann J, De Portu S, Smith-Palmer J, Roze S. Cost-effectiveness analysis of the MiniMed 670G hybrid closed-loop system versus continuous subcutaneous insulin infusion for treatment of type 1 diabetes. Diabetes Technol Ther. 2019;21:110–8.

52. Pease A, Zomer E, Liew D, Earnest A, Soldatos G, Ademi Z, et al. Cost-effectiveness analysis of a hybrid closed-loop system versus multiple daily injections and capillary glucose testing for adults with type 1 diabetes. Diabetes Technol Ther. 2020;22:812–21.

53. Pease A, Callander E, Zomer E, Abraham MB, Davis EA, Jones TW, et al. The cost of control: cost-effectiveness analysis of hybrid closed-loop therapy in youth. Diabetes Care. 2022;45:1971–80.

54. Jendle J, Buompensiere MI, Holm AL, de Portu S, Malkin SJP, Cohen O. The Cost-effectiveness of an advanced hybrid closed-loop system in people with type 1 diabetes: a health economic analysis in Sweden. Diabetes Ther. 2021;12:2977–91.

55. Serné EH, Roze S, Buompensiere MI, Valentine WJ, De Portu S, de Valk HW. Cost-effectiveness of hybrid closed loop insulin pumps versus multiple daily injections plus intermittently scanned glucose monitoring in people with type 1 diabetes in The Netherlands. Adv Ther. 2022;39:1844–56.

56. Wan W, Skandari MR, Minc A, Nathan AG, Winn A, Zarei P, et al. Cost-effectiveness of continuous glucose monitoring for adults with type 1 diabetes compared with self-monitoring of blood glucose: the DIAMOND Randomized Trial. Diabetes Care. 2018;41:1227–34.

57. García-Lorenzo B, Rivero-Santana A, Vallejo-Torres L, Castilla-Rodríguez I, García-Pérez S, García-Pérez L, et al. Cost-effectiveness analysis of real-time continuous monitoring glucose compared to self-monitoring of blood glucose for diabetes mellitus in Spain. J Eval Clin Pract. 2018;24:772–81.

58. Chaugule S, Graham C. Cost-effectiveness of G5 Mobile continuous glucose monitoring device compared to self-monitoring of blood glucose alone for people with type 1 diabetes from the Canadian societal perspective. J Med Econ. 2017;20:1128–35.

59. Roze S, Isitt J, Smith-Palmer J, Javanbakht M, Lynch P. Long-term Cost-Effectiveness of Dexcom G6 Real-time Continuous Glucose Monitoring Versus Self-Monitoring of Blood Glucose in Patients With Type 1 Diabetes in the U.K. Diabetes Care. 2020;43:2411–7.

60. Roze S, Isitt JJ, Smith-Palmer J, Lynch P. Evaluation of the long-term cost-effectiveness of the Dexcom G6 continuous glucose monitor versus self-monitoring of blood glucose in people with type 1 diabetes in Canada. Clinicoecon Outcomes Res. 2021;13:717–25.

61. Roze S, Isitt JJ, Smith-Palmer J, Lynch P, Klinkenbijl B, Zammit G, et al. Long-term cost-effectiveness the Dexcom G6 real-time continuous glucose monitoring system compared with self-monitoring of blood glucose in people with type 1 diabetes in France. Diabetes Ther. 2021;12:235–46.

62. Palmer AJ, Roze S, Valentine WJ, Minshall ME, Foos V, Lurati FM, et al. The CORE Diabetes Model: Projecting long-term clinical outcomes, costs and cost-effectiveness of interventions in diabetes mellitus (types 1 and 2) to support clinical and reimbursement decision-making. Curr Med Res Opin. 2004;20(Suppl):1.

63. Palmer AJ, Roze S, Valentine WJ, Minshall ME, Foos V, Lurati FM, et al. Validation of the CORE Diabetes Model against epidemiological and clinical studies. Curr Med Res Opin. 2004;20(Suppl):1.

64. McEwan P, Foos V, Palmer JL, Lamotte M, Lloyd A, Grant D. Validation of the IMS CORE Diabetes Model. Value Health. 2014;17:714–24.

65. Thokala P, Kruger J, Brennan A, Basarir H, Duenas A, Pandor A, et al. Assessing the cost-effectiveness of type 1 diabetes interventions: the Sheffield type 1 diabetes policy model. Diabet Med. 2014;31:477–86.

66. Pickup JC, Freeman SC, Sutton AJ. Glycaemic control in type 1 diabetes during real time continuous glucose monitoring compared with self monitoring of blood glucose: meta-analysis of randomised controlled trials using individual patient data. BMJ. 2011;343:d3805.

67. Ly TT, Nicholas JA, Retterath A, Lim EM, Davis EA, Jones TW. Effect of sensor-augmented insulin pump therapy and automated insulin suspension vs standard insulin pump therapy on hypoglycemia in patients with type 1 diabetes: a randomized clinical trial. JAMA. 2013;310:1240–7.

68. Beck RW, Riddlesworth T, Ruedy K, Ahmann A, Bergenstal R, Haller S, et al. Effect of continuous glucose monitoring on glycemic control in adults with type 1 diabetes using insulin injections: the DIAMOND Randomized Clinical Trial. JAMA. 2017;317:371–8.

69. Polonsky WH, Hessler D, Ruedy KJ, Beck RW. The impact of continuous glucose monitoring on markers of quality of life in adults with type 1 diabetes: further findings from the DIAMOND Randomized Clinical Trial. Diabetes Care. 2017;40:736–41.

70. Charleer S, De Block C, Van Huffel L, Broos B, Fieuws S, Nobels F, et al. Quality of life and glucose control after 1 year of nationwide reimbursement of intermittently scanned continuous glucose monitoring in adults living with type 1 diabetes (FUTURE): a prospective observational real-world cohort study. Diabetes Care. 2020;43:389–97.

71. Yeh H, Brown T, Maruthur N, Ranasinghe P, Berger Z, Suh Y, et al. Comparative effectiveness and safety of methods of insulin delivery and glucose monitoring for diabetes mellitus: a systematic review and meta-analysis. Ann Intern Med. 2012;157:390–400.

72. Nørgaard K, Scaramuzza A, Bratina N, Lalić NM, Jarosz-Chobot P, Kocsis G, et al. Routine sensor-augmented pump therapy in type 1 diabetes: the INTERPRET study. Diabetes Technol Ther. 2013;15:273–80.

73. Currie CJ, Morgan CL, Poole CD, Sharplin P, Lammert M, McEwan P. Multivariate models of health-related utility and the fear of hypoglycaemia in people with diabetes. Curr Med Res Opin. 2006;22:1523–34.

74. Gómez AM, Grizales AM, Veloza A, Marín A, Muñoz OM, Rondón MA. Factores asociados con el control glucémico óptimo en pacientes tratados con bomba de insulina y monitorización continua de glucosa en tiempo real. Avances en Diabetología. 2013;29:74–80.

75. Rios Y, García-Rodríguez J, Sanchez E, Alanis A, Ruiz-Velázquez E, Pardo A. Comparative effectiveness and safety of methods of insulin delivery and glucose monitoring for diabetes mellitus: a systematic review and meta-analysis. Ann Intern Med. 2012;157:390–400.

76. Pease A, Lo C, Earnest A, Kiriakova V, Liew D, Zoungas S. The efficacy of technology in type 1 diabetes: a systematic review, network meta-analysis, and narrative synthesis. Diabetes Technol Ther. 2020;22:411–21.

77. Swedisch National Diabetes Register. Annual report. 2013. https://www.ndr.nu/pdfs/Annual_Report_NDR_2013.pdf. Accessed 13 Oct 2023.

78. Nathan D, Cleary P, Backlund J, Genuth S, Lachin J, Orchard T, et al. Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. N Engl J Med. 2005;353:2643–53.

79. Bergenstal RM, Garg S, Weinzimer SA, Buckingham BA, Bode BW, Tamborlane WV, et al. Safety of a hybrid closed-loop insulin delivery system in patients with type 1 diabetes. JAMA. 2016;316:1407–8.

80. Collyns OJ, Meier RA, Betts ZL, Chan DSH, Frampton C, Frewen CM, et al. Improved glycemic outcomes with medtronic MiniMed advanced hybrid closed-loop delivery: results from a randomized crossover trial comparing automated insulin delivery with predictive low glucose suspend in people with type 1 diabetes. Diabetes Care. 2021;44:969–75.

81. Davis SN, Horton ES, Battelino T, Rubin RR, Schulman KA, Tamborlane WV. STAR 3 randomized controlled trial to compare sensor-augmented insulin pump therapy with multiple daily injections in the treatment of type 1 diabetes: research design, methods, and baseline characteristics of enrolled subjects. Diabetes Technol Ther. 2010;12:249–55.

82. Carlson AL, Sherr JL, Shulman DI, Garg SK, Pop-Busui R, Bode BW, et al. Safety and glycemic outcomes during the Mini-MedTM advanced hybrid closed-loop system pivotal trial in adolescents and adults with type 1 diabetes. Diabetes Technol Ther. 2022;24:178–89.

83. Carlson AL, Bode BW, Brazg RL, Christiansen MP, Garg SK, Kaiserman K, et al. 97-LB: safety and glycemic outcomes of the MiniMed Advanced Hybrid Closed-Loop (AHCL) system in subjects with T1D. Diabetes. 2020;69 (Supplement_1):97–LB.

84. Riveline JP, Schaepelynck P, Chaillous L, Renard E, Sola-Gazagnes A, Penfornis A, et al. Assessment of patient-led or physician-driven continuous glucose monitoring in patients with poorly controlled type 1 diabetes using basal-bolus insulin regimens: a 1-year multicenter study. Diabetes Care. 2012;35:965–71.

85. Lynch P, Attvall S, Persson S. Routine use of personal glucose monitoring system with insulin pump use in Sweden. Diabetologica. 2012;55:S432.

86. Garg SK, Weinzimer SA, Tamborlane WV, Buckingham BA, Bode BW, Bailey TS, et al. Glucose outcomes with the in-home use of a hybrid closed-loop insulin delivery system in adolescents and adults with type 1 diabetes. Diabetes Technol Ther. 2017;19:155–63.

87. Tang X, Yan X, Zhou H, Yang X, Niu X, Liu J, et al. Prevalence and identification of type 1 diabetes in Chinese adults with newly diagnosed diabetes. Diabetes Metab Syndr Obes. 2019;12:1527–41.

88. Zhou X, Wang X, An Y, Su Q, Li B, Chen H. Characteristics of type 1 diabetes patients aged 60 and older in Shanghai. J Endocr Disord. 2020;6:1039.

89. Bolinder J, Antuna R, Geelhoed-Duijvestijn P, Kröger J, Weitgasser R. Novel glucose-sensing technology and hypoglycaemia in type 1 diabetes: a multicentre, non-masked, randomised controlled trial. Lancet. 2016;388:2254–63.

90. Lind M, Polonsky W, Hirsch IB, Heise T, Bolinder J, Dahlqvist S, et al. Continuous glucose monitoring vs conventional therapy for glycemic control in adults with type 1 diabetes treated with multiple daily insulin injections: the GOLD Randomized Clinical Trial. JAMA. 2017;317:379–87.

91. Tumminia A, Crimi S, Sciacca L, Buscema M, Frittitta L, Squatrito S, et al. Efficacy of real-time continuous glucose monitoring on glycaemic control and glucose variability in type 1 diabetic patients treated with either insulin pumps or multiple insulin injection therapy: a randomized controlled crossover trial. Diabetes Metab Res Rev. 2015;31:61–8.

92. Tamborlane W, Beck R, Bode B, Buckingham B, Chase H, Clemons R, et al. Continuous glucose monitoring and intensive treatment of type 1 diabetes. N Engl J Med. 2008;359:1464–76.

93. Pease A, Earnest A, Ranasinha S, Nanayakkara N, Liew D, Wischer N, et al. Burden of cardiovascular risk factors and disease among patients with type 1 diabetes: results of the Australian National Diabetes Audit (ANDA). Cardiovasc Diabetol. 2018;17(1):77.

94. Thabit H, Tauschmann M, Allen JM, Leelarathna L, Hartnell S, Wilinska ME, et al. Home Use of an Artificial Beta Cell in Type 1 Diabetes. N Engl J Med. 2015;373:2129–40.

95. Benhamou PY, Franc S, Reznik Y, Thivolet C, Schaepelynck P, Renard E, et al. Closed-loop insulin delivery in adults with type 1 diabetes in real-life conditions: a 12-week multicentre, open-label randomised controlled crossover trial. Lancet Digit Health. 2019;1:e17-25.

96. Abraham MB, De Bock M, Smith GJ, Dart J, Fairchild JM, King BR, et al. Effect of a Hybrid Closed-Loop System on Glycemic and Psychosocial Outcomes in Children and Adolescents With Type 1 Diabetes: A Randomized Clinical Trial. JAMA Pediatr. 2021;175(12):1227–35.

97. Maiorino MI, Signoriello S, Maio A, Chiodini P, Bellastella G, Scappaticcio L, et al. Effects of continuous glucose monitoring on metrics of glycemic control in diabetes: a systematic review with meta-analysis of randomized controlled trials. Diabetes Care. 2020;43:1146–56.

98. Beaudet A, Clegg J, Thuresson PO, Lloyd A, McEwan P. Review of utility values for economic modeling in type 2 diabetes. Value Health. 2014;17:462–70.

99. Evans M, Khunti K, Mamdani M, Galbo-Jørgensen CB, Gundgaard J, Bøgelund M, et al. Health-related quality of life associated with daytime and nocturnal hypoglycaemic events: a time trade-off survey in five countries. Health Qual Life Outcomes. 2013;11:90.

100. McBride M, Eggleston AS, Jones T, Ly T. Health-Related Quality of life in patients with type 1 diabetes and impaired hypoglycaemia awareness: the role of sensor-augmented insulin pump therapy with automated insulin suspension. Value in Health. 2013;16:A448.

101. Palmer AJ, Si L, Tew M, Hua X, Willis MS, Asseburg C, et al. Computer modeling of diabetes and its transparency: a report on the eighth mount hood challenge. Value Health. 2018;21:724–31.

102. Kent S, Becker F, Feenstra T, Tran-Duy A, Schlackow I, Tew M, et al. The challenge of transparency and validation in health economic decision modelling: a view from mount hood. Pharmacoeconomics. 2019;37:1305–12.

103. Si L, Willis MS, Asseburg C, Nilsson A, Tew M, Clarke PM, et al. Evaluating the ability of economic models of diabetes to simulate new cardiovascular outcomes trials: a report on the ninth mount hood diabetes challenge. Value Health. 2020;23:1163–70.

104. Altunkaya J, Lee JS, Tsiachristas A, Waite F, Freeman D, Leal J. Appraisal of patient-level health economic models of severe mental illness: systematic review. Br J Psychiatry. 2021;220(2):1–12.

105. Silva-Illanes N, Espinoza M. Critical Analysis of Markov Models Used for the Economic Evaluation of Colorectal Cancer Screening: A Systematic Review. Value Health. 2018;21:858–73.

106. Teo E, Hassan N, Tam W, Koh S. Effectiveness of continuous glucose monitoring in maintaining glycaemic control among people with type 1 diabetes mellitus: a systematic review of randomised controlled trials and meta-analysis. Diabetologia. 2022;65:604–19.

Authors and Affiliations

Lisa A. de Jong[1] [ORCID] · **Xinyu Li[2]** [ORCID] · **Sajad Emamipour[3]** [ORCID] · **Sjoukje van der Werf[4]** [ORCID] · **Maarten J. Postma[1,5,6,7]** · **Peter R. van Dijk[8,10]** [ORCID] · **Talitha L. Feenstra[2,9]** [ORCID]

✉ Talitha L. Feenstra
t.l.feenstra@rug.nl

1 Department of Health Sciences, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

2   Groningen Research Institute of Pharmacy (GRIP), Faculty of Science and Engineering, University of Groningen, Groningen, The Netherlands

3   Department of Clinical Pharmacy and Pharmacology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

4   Central Medical Library, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

5   Department of Economics, Econometrics and Finance, Faculty of Economics and Business, University of Groningen, Groningen, The Netherlands

6   Center of Excellence for Pharmaceutical Care Innovation, Universitas Padjadjaran, Bandung, Indonesia

7   Department of Pharmacology and Therapy, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia

8   Department of Endocrinology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

9   National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

10 Department of Internal Medicine, Diabetes Center, Isala, Zwolle, The Netherlands