Review

A Comprehensive Review of Continuous Glucose Monitoring Accuracy during Exercise Periods

Elena Muñoz Fabra [1] , **José-Luis Díez** [1,2] , **Jorge Bondia** [1,2,*] and **Alejandro José Laguna Sanz** [2]

1 Instituto Universitario de Automática e Informática Industrial, Universitat Politècnica de València, Camino de Vera s/n, 46022 València, Spain; emunozfabra@gmail.com (E.M.F.); jldiez@isa.upv.es (J.-L.D.)
2 Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), Instituto de Salud Carlos III, 28029 Madrid, Spain; allasan@upvnet.upv.es
* Correspondence: jbondia@isa.upv.es

Abstract: Continuous Glucose Monitoring (CGM) has been a springboard of new diabetes management technologies such as integrated sensor-pump systems, the artificial pancreas, and more recently, smart pens. It also allows patients to make better informed decisions compared to a few measurements per day from a glucometer. However, CGM accuracy is reportedly affected during exercise periods, which can impact the effectiveness of CGM-based treatments. In this review, several studies that used CGM during exercise periods are scrutinized. An extensive literature review of clinical trials including exercise and CGM in type 1 diabetes was conducted. The gathered data were critically analysed, especially the Mean Absolute Relative Difference (MARD), as the main metric of glucose accuracy. Most papers did not provide accuracy metrics that differentiated between exercise and rest (non-exercise) periods, which hindered comparative data analysis. Nevertheless, the statistic results confirmed that CGM during exercise periods is less accurate.

Keywords: continuous glucose monitoring; type 1 diabetes; physical activity; accuracy; exercise; mean absolute relative difference

1. Introduction

Type 1 Diabetes (T1D) is a health condition in which insulin secretion by the pancreas is impaired or is completely missing, causing high levels of blood glucose in the affected patients. According to the International Diabetes Federation (IDF) [1], seven-point-seven billion people in the world population have diabetes. Of this figure, one-million, one-hundred-ten-thousand, one-hundred children and adolescents are known to suffer from T1D. In general, it can be estimated that around 5–10% of all diabetes cases correspond to T1D. This clearly has a great economic impact, such that by 2019, the IDF estimated that total diabetes-related health expenditure reached USD 760 billion.

Glucose homoeostasis is maintained naturally for individuals without diabetes, but people with T1D must control blood glucose concentration daily by means of exogenous insulin delivery. In recent years, several medical devices have been introduced to facilitate the management of insulin treatments for T1D. Continuous Glucose Monitoring (CGM) systems are some of the most popular, which provide continuous information about glucose levels based on a subcutaneously inserted probe that estimates glucose concentrations in blood based on interstitial electro-chemical measurements.

CGM has allowed real-time detection of changes in glucose levels and its immediate control. Even though the use of CGM is widespread among people with T1D, several studies showed [2–7] an important lack of accuracy under rapid and unexpected glucose rate changes, as occurs while doing exercise, commonly increasing the risk of hypoglycaemia. As these studies suggest, an understanding of this loss of accuracy during exercise might lead to compensation schemes mitigating its impact [8].

In this report, original studies that use CGM in T1D patients during different types of exercises are collected and compared. Specifically, the Mean Absolute Relative Difference (MARD) is analysed as a metric of the overall accuracy of CGM, and the MARD difference between the rest and exercise periods of the patients are calculated (ΔMARD), which provide a differential measurement of performance for exercise with respect to the "rest" periods. Even though there have been publications that review and compare studies that use CGM during exercise [9,10], these do not overlap with our paper, as they do not specifically analyse CGM accuracy according to the MARD value obtained, and they do not compare MARD values during exercise periods and the rest counterparts.

In this paper, we first describe the methods and exclusion criteria used for selecting the chosen articles (Section 2); next, the studies selected are screened (Section 3.1). Then, the results (Section 3) are discussed (Section 4). Finally, the conclusion reached, due to the given information, is presented (Section 5).

2. Materials and Methods

2.1. Eligibility Criteria

In this review, articles that reported original data on the use of CGM systems in T1D patients during a period of exercise were screened, selected and analysed. No distinction on the age of the patients was made.

The essential data extracted from each article were: publication year, cohort information (number of patients, age and sex), type(s) of exercise, total number of samples during exercise, glucose sensor and reference used for the computation of MARD. For those studies that provided an MARD value for the exercise periods, the total number of samples (obtained as the accumulated number of samples during exercise for all the trial participants) was extracted, and for those that also provided the MARD during rest periods, these data were also registered. Exercise intensity (% VO2 max) and duration (min) were also extracted for the analysis.

Discrepancies about the papers were resolved through consensus. Those articles that did not or not explicitly report the outcomes of interest were excluded.

Every study selected was classified, after critical scrutiny of the exercise description therein, into one of the next types:

- Aerobic: Patients usually exercised cycling (cycloergometer) or walking on a treadmill completing bouts during an amount of time and certain intensity of a fraction of the patients' maximum capacity. Rest time was also registered.
- Resistance: Patients exercised doing bouts of weight lifting exercises, ensuring that major muscles groups were targeted (leg press, bench press, leg curl, lat pull-down, abdominal crunches, shoulder press, seated row, etc.). Exercise was performed at a certain intensity and establishing the time between one bout and the next one.
- High Intensity Interval Exercise (HIIE): Patients exercised doing periods of maximum intensity exercise, usually involving very fast repetitions, alternated with short periods of resting or low-to-moderate intensity exercise. This type of exercise usually ends with exhaustion, and protocols can be very diverse.
- Intermittent: Patients exercised in intervals in which the intensity of the activity varied between different levels. Each interval is repeated cyclically until a determinate resting time is reached. For example, a one hour light running exercise with periodic bouts of faster sprint reaching 80% VO2 max every 10 min would qualify as intermittent.

A classification based on the method used to compute reference glucose for the computation of the MARD was also carried out:

- Gold standard: measurements were made in the laboratory by a glucose analyser from blood samples (Yellow Springs Instrument (YSI) or an equivalent device).
- Self-Monitoring Blood Glucose (SMBG): a glucometer was used involving fingerstick measurements.

MARD measurements for both the exercise and rest periods were the primary outcome extracted from the screened work to ensure that a proper statistical analysis of the data could be made and the accuracy of the CGM devices could be determined. When the MARD was not reported during the rest period, it was obtained (for the same sensor) from other articles, the main purpose of which was to study a determinate sensor accuracy without inclusion of exercise sessions [11,12].

2.2. Search Methods

An exhaustive electronic search was performed by using the National Center for Biotechnology Information (NCBI) library, called the National Library of Medicine (NLM) (including PubMed and PMC) [13]. The search strategy consisted of articles that included the keywords: CGM, continuous glucose monitoring, real time glucose monitoring, type 1 diabetes, exercise, physical activity, aerobic, resistance exercises, MARD. Once an article was selected and analysed, its references were examined, as long as they also included the same keywords. Furthermore, the recommendations provided by the web in the section "Similar Articles" were also revised. No language filter or date restrictions were applied.

2.3. Study Selection

Following the initial search, the title and abstract of the articles were analysed, and those that matched the inclusion criteria (Section 2.1) were further explored by thoroughly reviewing the nature of the studies described. Then, the content of the full document and the data confirmed its selection. Accuracy metrics were then extracted and stored for each of the selected studies.

2.4. Statistical Methods

Mean and SD values for the MARD during exercise and rest periods were extracted from the selected publications. Since clinical and statistical heterogeneity was expected among studies, a random effects model meta-analysis was conducted using inverse variance weighting for pooling [14–16]. The Hartung–Knapp method (also proposed by Sidik and Jonkman) was used for this purpose, since superior accuracy has been reported, especially when the number of studies is small [17,18]. The analysis was implemented in RStudio (RStudio Boston, MA, USA) using the function metacont, which was configured to use the standardised mean difference as the summary measure (Hedges' method [19]), the Sidik–Jonkman method for the estimation of between-study variance, and the computation of prediction intervals with the method proposed by Higgins et al. [20]. Furthermore, subgroup analysis was performed using the RStudio metamean function, which computes the overall mean from studies reporting a single mean using the inverse variance method for pooling. The function was configured to use the method described in Luo et al. [21] to estimate missing sample mean values from the reported sample size, median, range, and/or interquartile range. All statistical hypothesis were tested at a 95% confidence level ($\alpha = 0.05$).

3. Results

3.1. Studies Selected

The search produced several studies relevant to the aforementioned goals: a total of 54 sources. The data were extracted and are summarized in Table 1, in order to provide the reader with a snapshot view of the CGM-exercise literature.

Table 1. Reported data in chronological order of publication of the surveyed studies that carried out exercise trials. GS stands for Gold Standard glucose reference. ΔMARD $= MARD_e - MARD_r$ where $MARD_e$ is the MARD of the samples within exercise periods and $MARD_r$ is the counterpart for the samples during rest periods. Asterisks (*) indicate that $MARD_r$ was extracted from the articles mentioned in Table 2. The shaded rows correspond to those papers that reported necessary data to assess sensor accuracy during exercise, 17 in total. The labels that will later be used in the graphics to uniquely identify a study and branch have also been added to the MARD exercise column.

Source	Year	Patients	Exercise Samples	Exercise	Sensor Used	MARD Rest (%)	MARD Exercise (%)	ΔMARD	Reference
Iscoe et al. [7]	2006	5	–	Cycling	Guardian Real-Time	–	–	–	SMBG
Fayolle et al. [22]	2006	9	–	Cycling	GlucoDay	–	–	–	GS
Adolfsson et al. [23]	2008	12	–	Scuba diving	CGMS	–	–	–	GS
Riddell et al. [24]	2011	25	–	Sport Camps	Guardian Real-Time	–	–	–	SMBG
Adolfsson et al. [4]	2011	18 20 21	1135	Soccer Skiing Golf	CGMS	*18	24 (a) 27 (b) 19 (c)	6 9 1	SMBG
Herrington et al. [5]	2012	12	–	Cycling	Dexcom Seven Plus	–	–	–	GS
Yardley et al. [25]	2012	12	–	Aerobic and resistance	Medtronic Gold CGM	–	–	–	GS
Kumareswaran et al. [2]	2012	12	–	Walking	Freestyle Navigator	–	–	–	GS
Kumareswaran et al. [26]	2013	10	–	Walking	Freestyle Navigator	–	–	–	GS
Yardley et al. [27]	2013	12	–	Aerobic and resistance	Medtronic Gold CGM	–	–	–	GS
Radermecker et al. [28]	2013	10	–	Cycling	Guardian Real-Time	–	–	–	GS
Yousef et al. [29]	2014	12	–	Skydiving simulation	iPro2	–	–	–	SMBG
Campbell et al. [30]	2015	9	–	Running and simulate game-play activities	Medtronic Gold CGM	–	–	–	GS
Moser et al. [31]	2015	8	–	Aerobic	Guardian Real-Time	–	–	–	SMBG
van Dijk et al. [32]	2016	10	–	Walking	iPro2	–	–	–	SMBG
Moser et al. [33]	2016	7	489	Continuous cyclometer HIIE cyclometer	Guardian Real-Time	*15.8	18.76 (1.a) 19.63 (1.b)	2.96 3.83	SMBG
Bally et al. [34]	2016	10	108 100	Intermittent cycling Continuous cycling	Dexcom G4	*13	13.3 (a) 13.6 (b)	0.3 0.6	GS

Table 1. *Cont.*

Source	Year	Patients	Exercise Samples	Exercise	Sensor Used	MARD Rest (%)	MARD Exercise (%)	ΔMARD	Reference
Taleb et al. [35]	2016	17	431 425	Intermittent	Dexcom G4 Platinum Paradigm Veo (Enlite2)	13.77 12.38	22.53 (a) 20.44 (b)	8.76 8.06	GS
Moser et al. [33]	2016	7	–	Cycling	Guardian Real-Time	–	–	–	GS
McAuley et al. [36]	2016	14	–	Cycling	Paradigm Veo (Enlite2)	–	–	–	GS
Aberer et al. [37]	2017	12	462 540 502	Cycling	FreeStyle Libre Dexcom G4 Platinum MiniMed 640G (Enlite2)	*16.7 *13 *14.2	8.7 (a) 15.7 (b) 19.4 (c)	-8 2.7 5.2	GS
Gawrecki et al. [38]	2017	29	–	Walking	Guardian Real-Time	–	–	–	SMBG
Jayawardene et al. [39]	2017	12	48 48	HIIE cycling Cycling	MiniMed 670G (Sensor3)	*10.3	10.5 (a) 9.9 (b)	0.2 0.4	GS
Zaharieva et al. [40]	2017	12	96 96	Aerobic Resistance	iPro2	9.86 8.15	12 (1.a) 6.96 (1.b)	2.14 -1.19	SMBG
Reddy et al. [41]	2017	10	–	Aerobic and resistance	Dexcom G4 Platinum or G5	–	–	–	SMBG
Quirós et al. [42]	2018	5	–	Aerobic and resistance	Paradigm Veo (Enlite2)	–	–	–	GS
Larose et al. [43]	2018	22	–	Cycling	Dexcom G4 Platinum	–	–	–	SMBG
Giani et al. [6]	2018	17	136 136	Intermittent	FreeStyle Libre	16.6	12.5 (a) 15.4 (b)	-4.1 -1.2	GS SMBG
Aronson et al. [44]	2018	17	–	HIIE	Dexcom G4 Platinum	–	–	–	GS
Reddy et al. [45]	2018	10	–	Aerobic	Dexcom G4 Platinum	–	–	–	SMBG
Biagi et al. [46]	2018	6	108 86	Aerobic Resistance	Paradigm Veo (Enlite2)	9.5 15.5	16.5 (a) 16.8 (b)	7 1.3	GS
Abdulrahman et al. [47]	2018	4	–	Rugby training	Paradigm Veo (Enlite2)	–	–	–	SMBG
Castle et al. [48]	2018	20	–	Aerobic	Dexcom G5	–	–	–	SMBG
Moser et al. [49]	2018	10	228 140	Aerobic	iPro2 MiniMed 640G (Enlite2)	16.4 17.9	27 (2.a) 23.1(2.b)	10.6 5.2	SMBG

Table 1. *Cont.*

Source	Year	Patients	Exercise Samples	Exercise	Sensor Used	MARD Rest (%)	MARD Exercise (%)	ΔMARD	Reference
Steineck et al. [50]	2019	13	2660	Cycling	Dexcom G4 Platinum	*13	16.5 (a) 15.2 (b)	3.5 2.2	GS
Burckhardt et al. [51]	2019	14	–	Aerobic	Dexcom G5	–	–	–	SMBG
Forlenza et al. [52]	2019	12	–	Aerobic	Dexcom G4 (505)	–	–	–	SMBG
Larose et al. [53]	2019	22	–	Aerobic	Dexcom G4 Platinum	–	–	–	SMBG
Li et al. [3]	2019	17	192	HIIE	Dexcom G4 Platinum	10.4	17.8	7.4	GS
Zaharieva et al. [54]	2019	17	204	Aerobic	Dexcom G4(505) or G5	8	13 (2.a)	5	SMBG
Moser et al. [55]	2019	10	845	Cycling	Freestyle Libre	13.7	22 (3.a)	8.3	SMBG
Zaharieva et al. [56]	2019	12	–	Resistance	iPro2	–	–	–	SMBG
Eshghi et al. [57]	2019	12	–	Resistance	iPro2	–	–	–	SMBG
Steineck et al. [58]	2019	14	–	Cycling	Dexcom G4 Platinum	–	–	–	SMBG
Gawrecki et al. [59]	2019	16	–	Football	Guardian Connect	–	–	–	SMBG
Lee et al. [60]	2019	12	–	Cycling	FreeStyle Libre Pro	–	–	–	SMBG
Moser et al. [61]	2019	10	–	Cycling	FreeStyle Libre	–	–	–	SMBG
Scott et al. [62]	2019	14	–	Cycling	Dexcom G4 Platinum	–	–	–	SMBG
Moser et al. [63]	2019	14	470	Cycling	FreeStyle Libre	*16.7	29.8 (4.a)	13.1	SMBG
Scott et al. [64]	2019	14	–	HIIE and MICT	Dexcom G4 Platinum	–	–	–	SMBG
McCarthy et al. [65]	2020	16	–	Cycling	Dexcom G6	–	–	–	Unknown
Brockman et al. [66]	2020	23	–	Resistance	Medtronic Gold CGM iPro2	– –	– –	–	GS
Fokkert et al. [67]	2020	14	414 311	Mountain biking	Guardian Connect FreeStyle Libre	*10.9 *16.7	29 (a) 22 (b)	18.1 5.3	SMBG
Guillot et al. [68]	2020	24	96 96 80	Aerobic Resistance HIIE	Dexcom G6	9.5 9.1 16.8	13.3(a) 13 (b) 12.4 (c)	3.8 3.9 -4.4	SMBG

The oldest study (2011) that reported CGM accuracy is Adolfsson et al. [4], in which the reported accuracy metric, the MARD, ranged between 19% and 27%. A total of 24 adolescents were tested in three different sports (soccer, golf and floorball) using SMBG as the reference. The sensor used was the CGMS with MiniMed Solutions software Version 3.0.

A few years later, in 2016, Moser et al. [33] presented a study that also used the SMBG method for glucose reference. Eight patients were tested using a cycle ergometer executing continuous and High Intensity Interval Exercise (HIIE) protocols, using the Guardian Real-Time. The MARD exercise remained in the same relative range (18–20%). In the same year, Bally et al. [34] also presented the MARD exercise within a good range (13.3% for intermittent and 13.6% for aerobic). They used the Dexcom G4 Platinum with YSI 2300 STAT Plus Glucose Analyser to calculate sensor accuracy. Jayawardene et al. [39] involved 12 patients in a trial that explored the impact of aerobic exercise (cycling), reporting a 9.9% MARD; and HIIE exercise with 10.5%, using also a YSI Analyser, but with the Medtronic Minimed 670G, which integrates the Guardian Sensor 3 as the sensing unit. Although this review focuses on sensing accuracy, the name of the integrated sensor-pump systems will still be used throughout the manuscript to keep faithful to the device description in the source papers. However, a codification will be added to highlight the sensor incorporated. Thus, this device will be referred to as Medtronic Minimed 670G-S3.

One of the most recent studies selected in this review is Fokkert et al. [67]. Fourteen patients were monitored during six days of mountain biking activity, using the Guardian-Connect and the FreeStyle Libre sensor, reporting a higher MARD than other studies for the exercise period, with values of 29% and 22%, respectively. On the other hand, Guillot et al. [68], which is also a recent study, presented a lower MARD for the three types of exercise, using the Dexcom G6 and SMBG to calculate the CGM accuracy.

Moser et al. [63] (2020) explored the FreeStyle Libre sensor accuracy on a cohort of 14 patients. The reported MARD value during exercise was higher than those previously mentioned: 29.8%. The MARD during the rest period was significantly lower, at 8.6%. This is one of the few studies found that reported both MARD values, which enables the possibility of a more detailed analysis of the impact of the MARD on exercise, since CGM accuracy can be normalized with respect to its rest counterpart, i.e., a ΔMARD can be computed. That is also the case for the study by Taleb et al. [35], which is also interesting since it studied the accuracy of two different sensors: Dexcom G4 Platinum, the data of which will be referenced as Taleb Dexcom in this analysis, and Paradigm Veo (Enlite 2), which will be referenced as Taleb Enlite. The integrated system composed of the Paradigm Veo and the Enlite 2 sensor, from now on, will be referred to as Paradigm Veo-E2. Both CGM systems were used in the same conditions by 17 patients. Aberer et al. [37] also studied three different sensors: FreeStyle Libre, Dexcom G4 Platinum and Medtronic MiniMed 640G (integrating the Enlite 2 sensor), but they only provided accuracy values for the exercise periods. This device will be referred to as Medtronic Minimed 640G-E2. Moser et al. [49] also tested multiple devices, recruiting 10 participants for aerobic exercises, using the iPro2 and MiniMed 640G-E2 in the same conditions. Steineck et al. [50] compared two different placements of Dexcom G4 on either the abdomen or the arm.

Giani et al. [6] explored the behaviours of 17 patients during intermittent exercise using the FreeStyle Libre, but with the particularity of reporting the MARD for both SMBG and YSI Analyser. It must be noted that the data from this study yielded negative ΔMARD values, which seems to contradict the general behaviour of CGM during exercise. This particularity will be further discussed in Section 4. Only one other study presented this issue, the result of resistance exercise in Zaharieva et al. [40]. From this article, it is worth noting the small standard deviations reported, ranging between 0.06 and 0.12 mg/dL. In both cases, twelve patients were monitored using the Paradigm Veo-E2 system with iPro2.

In Li et al. [3], six patients were monitored using the Paradigm Veo-E2, and it was one of the few studies that implemented HIIE as the exercise protocol, using Dexcom G4 Platinum to monitor 17 patients. The reported value for exercise was 17.8%, and ΔMARD

was 7.4%, quite similar to Biagi et al.'s [46] data during aerobic exercise. However, the data from the same study [46] during resistance exercise reported a much lower ΔMARD.

In most of the mentioned articles, the CGM devices were calibrated using SMBG. Those that used FreeStyle Libre were already factory-calibrated, for example Giani et al. [6] and Aberer et al. [37]. Some articles did not specify the calibration method used, like Moser et al. [55], Williams et al. [63], Fokkert et al. [67] and Guillot et al. [68].

For the sake of completion, accuracy data for sensors used in studies that did not provide the MARD for rest periods were searched in complementary articles. An example of this is Mastrototaro et al. [69], which tested the accuracy and efficacy of the Guardian Real-Time CGM in 72 subjects, obtaining retrospectively from an open-label, multicenter, six month study. Slover et al. [70] used a fourth-generation sensor, the Guardian Connect, tested in 145 patients, and Rodbard et al [71] summarized some of the current CGM systems from Medtronic, including the integrated system MiniMed 640G with the Enlite sensor. Nakamura et al. [72] monitored 72 subjects who were enrolled at four US centres, wore the Dexcom G4 Platinum for up to seven days, and participated in a total of 36 h of monitoring, using YSI Analyser. Garg et al. [73] studied 30 adolescents and 94 adults using MiniMed 670G-S3 with a reference measurement by an i-STAT device. Gross et al. [74] monitored the glucose of 135 patients from eight clinical sites using the CGMS during patient home use for three days or more.

Complementary articles were matched only by the type of sensor used. Other factors that could affect sensor accuracy, such as age [75] or protocol characteristics, were unfeasible to be matched, and furthermore, it was desired to maintain a simple relationship between paired MARD values. However, to avoid confounding factors, only studies reporting both MARD during exercise and rest periods were included in the ΔMARD meta-analysis study. Matched articles are marked with an asterisk in Table 1. These data are organised in Table 2.

Table 2. Reported MARD data for resting periods for the sensors that the studies in Table 1 did not originally provide.

| Paper Extracted	Sensor Used	MARD Rest (%)	Reference
Gross et al. [74]	CGMS	18	SMBG
Mastrototaro et al. [69]	Guardian Real-Time	15.8	SMBG
Nakamura et al. [72]	Dexcom G4 Platinum	13	GS
Rodbard et al. [71]	MiniMed 640G-E2	14.2	GS
Garg et al. [73]	MiniMed 670G-S3	10.3	GS
Hansen et al. [11]	FreeStyle Libre System	16.7	SMBG
Slover et al. [70]	Guardian-Connect	10.9	SMBG

The article of Gross et al. [74], which studied the CGMS accuracy, provided the MARD during rest periods for Adolfsson et al. [4], 18%. In the case of Moser et al. [33], this value was assigned by Mastrototaro et al. [69], which was 15.8%. Nakamura et al. [72] provided the MARD during rest periods for the Dexcom G4, 13%, used in Bally et al. [34], Aberer et al. [37] and Steineck et al. [50]. Garg et al. [73] provided a value of 10.3% for the MARD during rest periods of the MiniMed 670G-S3, which was compatible with the sensor used in Jayawardene et al. [39]. The value for the FreeStyle Libre System is needed to complement the studies by Aberer et al. [37], Moser et al. [63] and Fokkert et al. [67]. According to Hansen et al. [11], this value is 16.7%. Slover et al. [70] gave the MARD during rest periods for the Guardian-Connect, 10.9%, which can be used to complement the study by Fokkert et al. [67]. Finally, the MiniMed 640G-E2 value, needed for Aberer et al. [37], was provided by Rodbard et al. [71], 14.2%.

In summary, from the original 54 sources listed in Table 1, sixteen provided MARD data during exercise. These sources are highlighted in Table 1 as shaded rows. These

MARD values are displayed in Figure 1 grouped by glucose reference (Figure 1a) and type of exercise (Figure 1b) in order to facilitate visual inspection of the data.

Even though the CGM accuracy improved overall with time, it is not trivial to conclude the same about the CGM accuracy during exercise periods. In Figure 2, data are displayed in chronological order in order to better visualize the technological advancements over time.
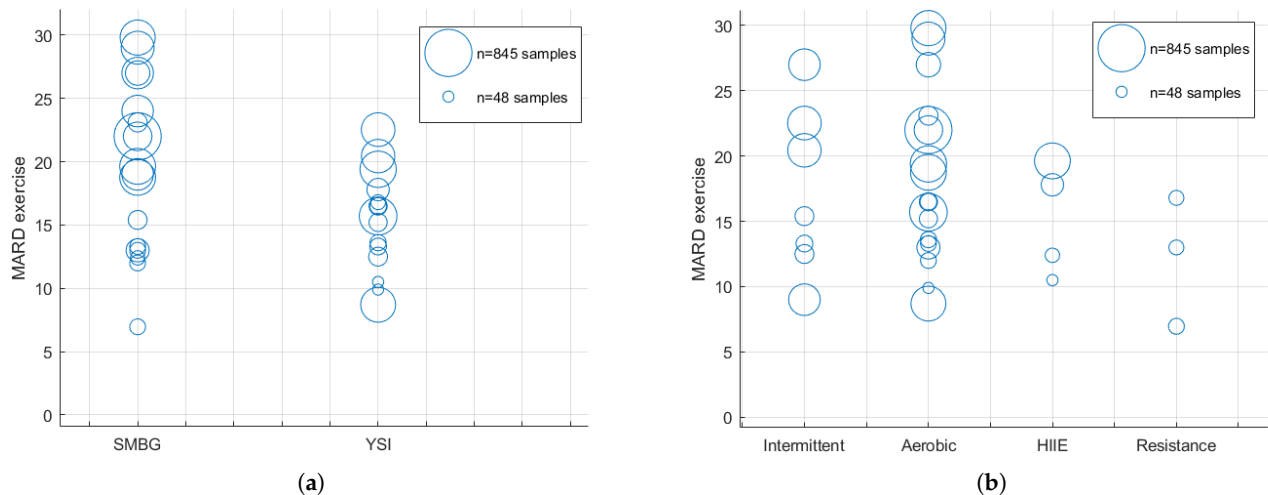
Figure 1. Panel (**a**) shows the MARD for the exercise periods according to the reference used. The SMBG column displays data from studies where self-monitoring blood glucose samples were used to calculate the CGM accuracy. The YSI column stands for those studies that used gold-standard reference methods (YSI or equivalent) to calculate the CGM accuracy. Panel (**b**) displays MARD data for the exercise periods grouped according to the type of exercise. The radius of each bubble is proportional to the number of samples used for the computation of MARD values.
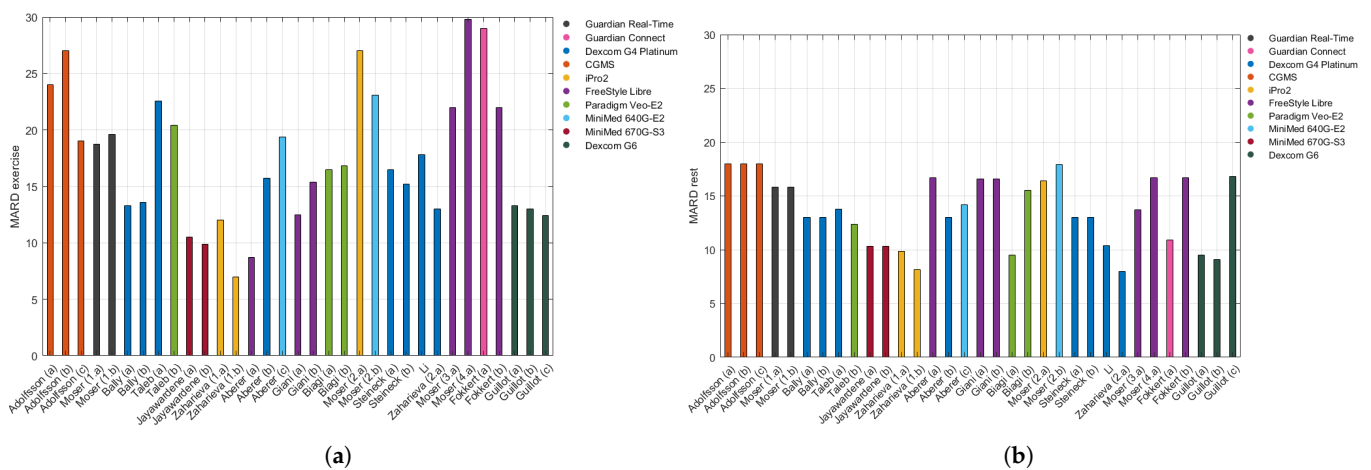
Figure 2. Panel (**a**) shows the MARD for the exercise periods in chronological order of the publication, coloured according to the CGM system used, and Panel (**b**) depicts the MARD of the rest periods, presented in the same way.

Figure 2 shows the MARD for exercise periods coloured according to the sensor used in each study:

- CGMS: Medtronic MiniMed Inc., Northridge, CA, USA.
- Guardian Real-Time: Medtronic MiniMed Inc., Northridge, CA, USA.
- Guardian Connect: Medtronic MiniMed Inc., Northridge, CA, USA.
- Dexcom G4 Platinum: Dexcom Inc., San Diego, CA, USA.
- iPro2: Medtronic MiniMed Inc., Northridge, CA, USA.
- FreeStyle Libre System: Abbott Diabetes Care, Maidenhead, UK.

- Paradigm Veo-E2: Medtronic MiniMed Inc., Northridge, CA, USA.
- MiniMed 640G-E2: Medtronic MiniMed Inc., Northridge, CA, USA.
- MiniMed 670G-S3: Medtronic MiniMed Inc., Northridge, CA, USA.
- Dexcom G6: Dexcom Inc., San Diego, CA, USA.

Assuming a relationship exists between CGM accuracy data for the rest periods from Table 2 and those originally listed in Table 1, it was possible to represent MARD data in Figure 2b as an analogous figure to Figure 2a.

However, only nine studies originally provided both the MARD during exercise and rest period. For those studies, ΔMARD was calculated to provide a relative measurement of accuracy caused only by the effect of the exercise, as seen in Table 1. In order to keep the flow of information similar to what has already been shown, the data were organized into two different graphs, with the colour palette depending on the reference used (Figure 3a) or the type of exercise (Figure 3b). Positive ΔMARD values indicate a larger error for a particular study during exercise compared to the baseline estimation error of that device and trial.

In order to aggregate the ΔMARD for data from Table 2 and those studies from Table 1 that did not provide MARD for the rest periods, the data are gathered into two new plots: according to the glucose reference used (Figure 3c) and the type of exercise (Figure 3d).

Figure 3a,b is analogous to Figure 3c,d, respectively, in which they are grouped according to the same factor. However, it was preferred not to represent them together in a unique graph, since it is important to remark that the MARD for the rest periods was obtained from different sources and may include a different bias from the MARD during exercise.

3.2. Meta-Analysis

The data used in the analysis are the MARD, the number of observations and the Standard Deviation (SD) during the exercise and rest period. In those articles for which the rest data were obtained from external papers (Table 2), the same data were assumed by different control groups. Therefore, in order to avoid "double-counting" the participants in the control group, the pooled effect size of those data was synthesized.

Some articles did not provide the SD, but provided the Interquartile Range (IQR), so in order to make them comparable to the rest, the formula proposed in the Cochrane Handbook [76] and Wan et al. [77] was used: $SD = (Q_3 - Q_1)/1.35$, where Q_3 and Q_1 stand for the third and first quartiles (75 and 25%). Adolfsson et al. [4] and Fokkert et al. [67] did not provide the SD or IQR of the CGM accuracy; thus, they were excluded from the analysis.

There were two articles that did not provide the SD or IQR of the CGM accuracy, Adolfsson et al. [4] and Fokkert et al. [67]; thus, they had to be excluded from the analysis.
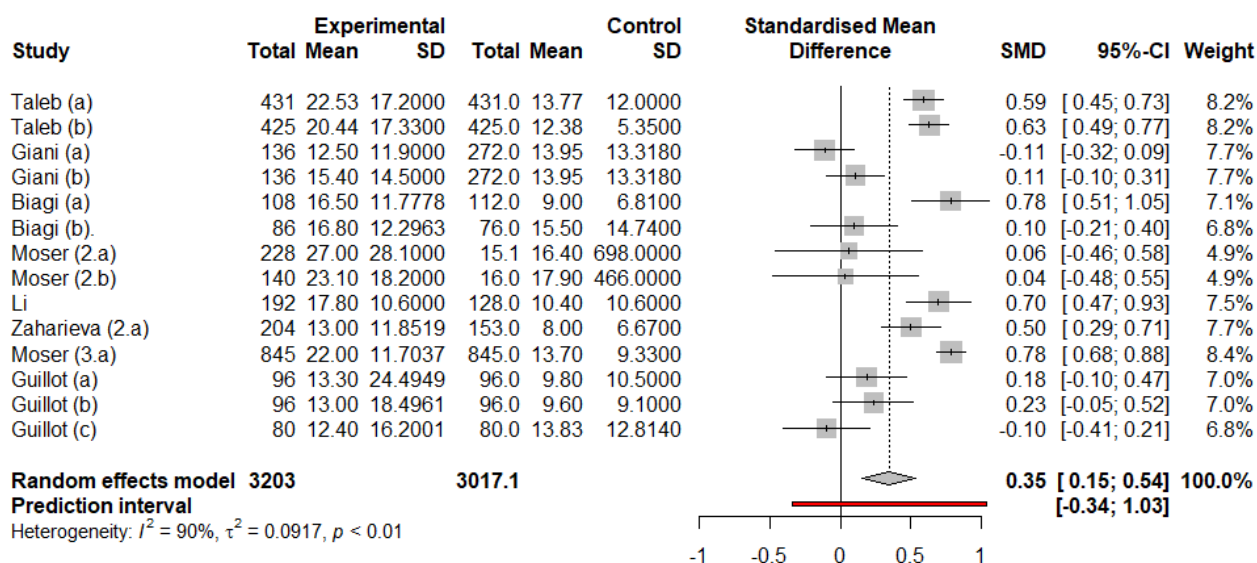
Another method [76] for obtaining the SD is to use the Confidence Interval (CI) values provided: $SD = \sqrt{N} \cdot (CImax - CImin)/3.92$, N being the number of samples. This method was applied for the case of Guillot et al. [68]. In that same article, only the median value for the absolute relative difference of the rest samples was reported. Thus, it was necessary to calculate the mean by applying the approximation proposed at Wan et al. [77].

A subgroup meta-analysis, taking into account the data from Tables 1 and 2 to populate with enough data each analysed subgroup, was done to factor the influence of the glucose reference in the exercise effect on the MARD, obtaining a *p*-value $p_\alpha = 0.1023$. Figure 1b shows the MARD value grouped by type of exercise, which could also be a differentiating factor that influenced the MARD during exercise (aerobic/resistance/intermittent/HIIE). Taking that into account, a new subgroup meta-analysis was performed, but now factoring the type of exercise. A *p*-value $p_\beta = 0.5729$ was calculated.

(a)

(b)

(c)

(d)
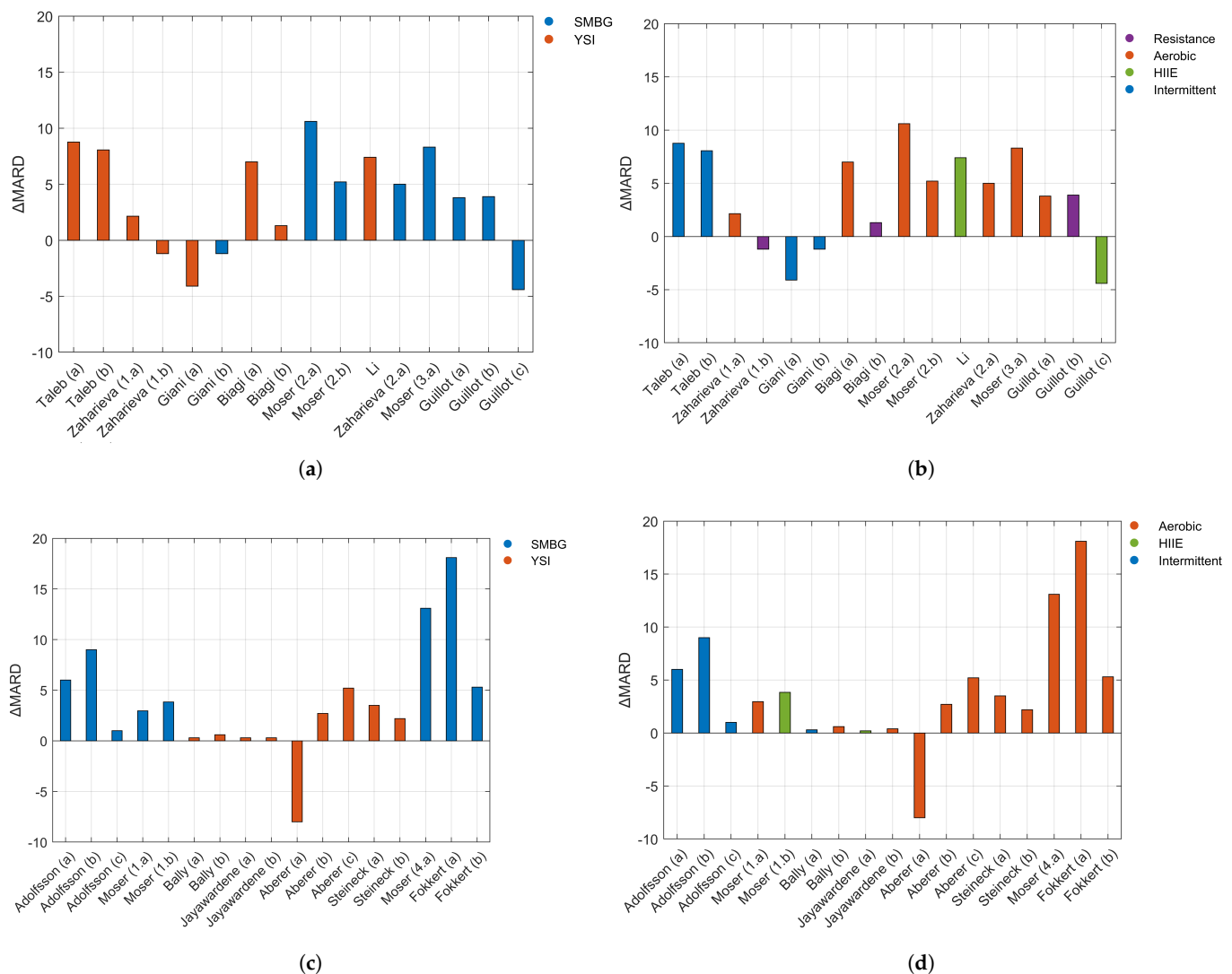
Figure 3. Panels (**a**,**b**) display data from studies that reported the MARD for both exercise and rest. Panel (**a**) shows ΔMARD according to the reference used, and Panel (**b**) shows the same data coloured according to the type of exercise. Panels (**c**,**d**) display data from studies that reported MARD values for exercise periods, but did not provide MARD values for resting periods. In order to calculate ΔMARD for those studies, MARD values for exercise were paired with the MARD from the sources in Table 2. Panel (**c**) displays ΔMARD according to the reference used, and Panel (**d**) presents the same data coloured according to the type of exercise. In Panels (**a**,**c**), YSI stands for those studies that used gold-standard reference methods (YSI or equivalent) to calculate the CGM accuracy.

Considering simultaneously the accuracy data from both exercise and rest periods as extracted in Table 1, without taking into account data from Table 2, a relative metric for sensor accuracy was calculated in ΔMARD, obtaining $p_\gamma = 0.8244$ for the estimation of its mean value. Data from Zaharieva et al. [40] were found to be quantitatively different than the rest of the data analysed; the reported standard deviation for that study was an order of magnitude lower than that of any other analysed papers, probably due to the retrospective calibration method of estimation of the iPro2 device. This in turn could affect the standardised mean difference (main metric for the meta-analysis), making it the highest and lowest values of all those in the analysis, greatly influencing in return its outcome. Those numbers were extreme and were considered outliers, and therefore excluded from the analysis. The resulting analysis after removing those data is shown in Figure 4, with a *p*-value $p_\delta = 0.0018$.

Study	Total	Experimental Mean	SD	Total	Control Mean	SD	Standardised Mean Difference	SMD	95%-CI	Weight
Taleb (a)	431	22.53	17.2000	431.0	13.77	12.0000		0.59	[0.45; 0.73]	8.2%
Taleb (b)	425	20.44	17.3300	425.0	12.38	5.3500		0.63	[0.49; 0.77]	8.2%
Giani (a)	136	12.50	11.9000	272.0	13.95	13.3180		-0.11	[-0.32; 0.09]	7.7%
Giani (b)	136	15.40	14.5000	272.0	13.95	13.3180		0.11	[-0.10; 0.31]	7.7%
Biagi (a)	108	16.50	11.7778	112.0	9.00	6.8100		0.78	[0.51; 1.05]	7.1%
Biagi (b).	86	16.80	12.2963	76.0	15.50	14.7400		0.10	[-0.21; 0.40]	6.8%
Moser (2.a)	228	27.00	28.1000	15.1	16.40	698.0000		0.06	[-0.46; 0.58]	4.9%
Moser (2.b)	140	23.10	18.2000	16.0	17.90	466.0000		0.04	[-0.48; 0.55]	4.9%
Li	192	17.80	10.6000	128.0	10.40	10.6000		0.70	[0.47; 0.93]	7.5%
Zaharieva (2.a)	204	13.00	11.8519	153.0	8.00	6.6700		0.50	[0.29; 0.71]	7.7%
Moser (3.a)	845	22.00	11.7037	845.0	13.70	9.3300		0.78	[0.68; 0.88]	8.4%
Guillot (a)	96	13.30	24.4949	96.0	9.80	10.5000		0.18	[-0.10; 0.47]	7.0%
Guillot (b)	96	13.00	18.4961	96.0	9.60	9.1000		0.23	[-0.05; 0.52]	7.0%
Guillot (c)	80	12.40	16.2001	80.0	13.83	12.8140		-0.10	[-0.41; 0.21]	6.8%
Random effects model	**3203**			**3017.1**				**0.35**	**[0.15; 0.54]**	**100.0%**
Prediction interval									**[-0.34; 1.03]**	

Heterogeneity: $I^2 = 90\%$, $\tau^2 = 0.0917$, $p < 0.01$

Figure 4. Meta-analysis using the random effects model, including the corresponding forest plot. The above-mentioned outliers were excluded. SMD stands for Standardised Mean Difference.

4. Discussion

Selecting the variables for the grouping and visualization of the MARD data was challenging. The fact that many studies have a similar length and intensity produced an overlap of most studies in these factors, rendering those studies impossible to separate.

The CGM device is one of the more relevant factors considered, as it changes among the different studies and years of literature. As technology evolves, the MARD is expected to be reduced with newer generations of CGM being introduced to the market. Thus, in general terms, the more recent sensor used, the better the technology and more accurate the CGM devices are. As may be observed in Figure 2a (the MARD exercise according to the sensor used), that trend can be appreciated: the oldest Medtronic sensors, CGMS (2000–2002) and Guardian Real-Time (2005), presented more error than the one integrated into the Minimed 670G-S3 (2016). The FreeStyle Libre, a Flash Glucose Monitoring (FGM) device requiring manual scanning, usually provided less accurate measurements [55,78].

There seems to be a mismatch between the marketed date of the CGM devices and their use in the studies screened. For example, Li et al. [3] (2019) and Zaharieva et al. [54] (2019) used the Dexcom G4 Platinum (2014). Even further, some studies used devices of more than 10 years old. For example, the Guardian-Real Time was used in Iscoe et al. [7] from 2006 and in the study of 2017 of Gawreick et al. [38]. It is likely that this date mismatch affected the analysis of the CGM accuracy over time. Perhaps the use of an obsolete sensor would not be the reason for less accurate results, but its influence cannot be ruled out either.

It is remarkable to find a lack of new-generation sensors used in exercise studies, such as Dexcom G5 or G6, even in more recent studies, when they present improved accuracy compared to previous generations according to the manufacturers. Zaharieva et al. [54] and Guillot et al. [68] are the only studies that provided MARD values using Dexcom G5 (MARD exercise 13%) and Dexcom G6 (MARD exercise 13.3%, 13% and 12.4%), respectively.

Regarding MiniMed 640G-E2, it was not possible to draw any conclusions, as it was used in only two papers. Other devices, like Medtronic 670G-S3, Dexcom G4 Platinum and Paradigm Veo-E2, presented lower errors. No clear relationship could be established between the CGM accuracy of these devices and the analysed factors.

The iPro2 initially seemed to be the most accurate sensor for exercise periods. It seems to be better than every other CGM system under study. The limitation with iPro2 CGM,

however, is that glucose values are not reported in real-time, since they use a retrospective calibration algorithm, and as such, it can only be used for retrospective analysis such as the one presented in [40].

At first glance when analysing Figure 1a, which shows the MARD under exercise classified per reference glucose used, studies using SMBG seem to have a greater MARD than those using YSI. To clarify this hypothesis, a meta-analysis was done on the data available. The p-value p_α, which was obtained from the subgroup meta-analysis between the studies that used SMBG and YSI, shows that the type of reference could not be determined as a relevant factor to the difference in the MARD between groups, as it was higher than 0.05. The type of exercise factor also resulted in being non-conclusive, as a p-value p_β of 0.5729, also greater than 0.05, was obtained from its subgroup meta-analysis.

The fact that a non-significant p-value p_γ was obtained for the initial hypothesis that the exercise is detrimental to the MARD does not allow the above-mentioned hypothesis to be corroborated and is contrary to what was previously observed. This led to a new meta-analysis displayed in the forest plot in Figure 4, which resulted in a statistically significant p-value $p_\delta = 0.0018$, confirming that CGM during exercise periods is less accurate. This has important implications for diabetes management, since people with diabetes must be aware about the possible inaccuracy of glucose sensing devices during exercise (i.e., periods of increased glucose fluctuations).

On the other hand, the analysis of the MARD in Figure 3a,b ΔMARD according to the reference use and type of exercise) provided interesting conclusions. It can be appreciated that the accuracy was lower during intermittent and aerobic exercises. Only two studied HIIE, thus no tendency can be determined. One of the cases presented a negative ΔMARD value, which means that the MARD for exercise was lower (12.4%) than the pre-exercise value (16.8%). With respect to the resistance exercises, a greater accuracy of the sensors can be observed, as it was concluded in individual studies that compared resistance exercise and other exercises [8,46]. One of the cases also presented a negative ΔMARD value, the MARD for exercise being 6.96% and the MARD for rest being 8.15%. This case did not alter the trend observed in the other studies.

Comparing the graphs of ΔMARD according to the reference used, in Figure 3a,c (left top and bottom graphs), it is worth noting that negative values of ΔMARD occur mostly when the reference is SMBG. It is also apparent that positive values for SMBG in both figures range between 2% and 9%. Regarding the graphs of ΔMARD according to the type of exercise, in Figure 3b,d, negative values do not coincide with the same type of exercise: in the upper graph, they correspond to intermittent and resistance activity, and in the bottom graphs to aerobic. In both figures, the positive values for intermittent exercise lay between 6% and 9%. Aerobic exercise in Figure 3d seems to show higher values, reaching a maximum of 18.1%, while in Figure 3b, the highest value is 8.3%.

For individuals with type 1 diabetes, aerobic exercise typically leads to the greatest risk of hypoglycaemia. CGM accuracy during exercise periods is therefore critical for detecting and potentially treating these situations. Patients need to know that they can rely on CGM systems during any kind of situation with a fast and accurate response. Further work needs to be done exploring the impact of CGM accuracy on aerobic exercise, and possibly correcting that impact, as proposed by Laguna et al. [8].

It is particularly interesting to consider the implications of the CGM accuracy on closed-loop artificial pancreas systems, which are the technological vanguard in T1D treatments. In Huyett et al. [79], a comparison of glucose sensing dynamics vs. closed loop performance was performed, showing that a lower sensor lag (and in turn, a lower MARD) yields better treatment by a closed-loop algorithm after meal ingestion. It is thus paramount for the adequate treatment of people with diabetes to ensure the accuracy of CGM during every stage of their lives, particularly during periods of great glycemic variability. This is especially critical during periods of high hypoglycaemic risk such as exercise bouts, as stated above.

In the literature analysed in this review, it was hard to establish a clear relationship between MARD and any factors relevant to either the exercise or the underlying diabetes condition. Despite the amount of studies that complied with the selection conditions, only some of them provided the MARD during exercise, and even fewer reported the MARD during rest periods. In other cases, the MARD was not provided as a mean value, but instead, the median ARD [2,26,53] or Mean Absolute Difference (MAD) [23,80] was reported. The main reason could be that the studies did not focus on the accuracy of the CGM during the exercise. Some of them were focused on the analysis of the heart rate in the different exercises and post-exercise periods [31,59,81], time spent in hypoglycaemia [42,60,65] or glucose levels [41,51,66].

Therefore, this work encourages further detail of the CGM accuracy in exercise periods in future studies, facilitating the proposal of a solution to solve CGM errors during periods with high risk of hypoglycaemia. Given that the blood glucose reference was not found to relevantly influence the effect of exercise on the accuracy of the sensors evaluated, we acknowledge the clear benefits of favouring SMBG devices (which keep improving in accuracy [82]) over gold-standard reference measurements, which are limited to laboratory analysis and IV blood sampling. It must be remarked that CGM devices are calibrated using SMBG, and accuracy cannot be expected to be better than that of the calibration device used. Furthermore, SMBG allows for more flexible outpatient exercise monitoring protocols, which tend to provide richer sets of data over more diverse cohorts of patients. Additionally, SMBG also facilitates gathering accuracy data from resting ambulatory periods at lower costs than those in a clinic environment. In addition, many of the in-clinic studies analysed are often preceded by days of free living conditions of the patients after CGM insertion. CGM data from these days, coupled with several daily SMBG measurements, could constitute the baseline MARD that is missing in many of the works that report the MARD for exercise, without increasing the cost of the experiments.

However, for those studies in which the primary goal is discerning the influence of exercise on the CGM accuracy, more detailed guidelines should be discussed. The studies analysed in this work reported widely different sample sizes with MARD values during rest periods and those with exercise. These differences in sample sizes could hinder the comparison of MARD values and the computation of the exercise effect. As recommended by Danne et al. [83], head-to-head studies are encouraged to avoid sample mismatches in the MARD estimations. As for the number of samples to collect, Reiterer et al. [84] stated that confidence in the estimation of the CGM accuracy greatly depends on the sample size of the study. Thus, ideally, the sample size must be maximized. In practice, it is often unreasonable to ask more for than 60 min of exercise in a trial from a patient (although it depends on the cohort). Additionally, if the MARD were to be monitored using SMBG, the patient would be asked to use a finger strip every several minutes, which is also unrealistic for long periods of exercise. Therefore, a compromise must be achieved: we propose that, for MARD characterization purposes, more than 45 min of exercise for each visit are scheduled, and at least four reference glucose samples are collected. The trial number of samples can be maximized by repeating the exercise visits with the same or different patients.

In summary, the main findings of this work are threefold: (1) exercise negatively affects CGM accuracy; (2) no clear statistical influence on the CGM accuracy was found for factors such as type of exercise or glucose reference used; and (3) few studies simultaneously reported CGM accuracy for both exercise and rest periods, which makes the analysis of the influence of exercise on CGM errors difficult.

5. Conclusions

In this review, an analysis of published articles that study people with T1D using CGM during exercise was performed. The results of the meta-analysis indicate that the accuracy of the monitoring devices is negatively affected by the exercise periods. This is of particular interest to know for both researchers, who could take it into consideration when

designing new exercise experiments, and patients or clinicians who try to manage diabetes, who must consider the presence of larger CGM errors during periods of exercise.

MARD data from multiple studies were pooled, visualized and statistically compared. It was shown that no clear statistical difference can be found in the precision of sensors for factors such as the blood glucose measurement method or the type of exercise. More modern sensors are expected to be more accurate during periods of exercise; however, it was found that very few studies for exercise have been done using the latest generation of sensors, to our surprise.

Author Contributions: Conceptualization, A.J.L.S. and E.M.F.; methodology, E.M.F., A.J.L.S., J.B. and J.-L.D.; software, E.M.F.; validation, A.J.L.S., J.B. and J.-L.D.; formal analysis, E.M.F. and A.J.L.S.; investigation, E.M.F.; resources, E.M.F. and A.J.L.S.; data curation, E.M.F. and A.J.L.S.; writing, original draft preparation, E.M.F.; writing, review and editing, A.J.L.S., J.B. and J.-L.D.; visualization, E.M.F. and A.J.L.S.; supervision, A.J.L.S., J.B. and J.-L.D.; project administration, A.J.L.S. and J.B.; funding acquisition, A.J.L.S. and J.B. All authors read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

T1D	Type 1 Diabetes
CGM	Continuous Glucose Monitoring
IDF	International Diabetes Federation
FGM	Flash Glucose Monitoring
MARD	Mean Absolute Relative Difference
MAD	Mean Absolute Difference
HIIE	High Intensity Interval Exercise
MICT	Moderate-Intensity Continuous Training
SMBG	Self-Monitoring Blood Glucose
NCBI	National Center for Biotechnology Information
NLM	National Library of Medicine
CG	Capillary Glucose
GS	Gold Standard

References

1. International Diabetes Federation. *IDF Diabetes Atlas*; IDF: Brussels, Belgium, 2019.
2. Kumareswaran, K.; Elleri, D.; Allen, J.M.; Caldwell, K.; Nodale, M.; Wilinska, M.E.; Amiel, S.A.; Hovorka, R.; Murphy, H.R. Accuracy of continuous glucose monitoring during exercise in type 1 diabetes pregnancy. *Diabetes Technol. Ther.* **2013**, *15*, 223–229. [CrossRef] [PubMed]
3. Li, A.; Riddell, M.C.; Potashner, D.; Brown, R.E.; Aronson, R. Time Lag and Accuracy of Continuous Glucose Monitoring during High Intensity Interval Training in Adults with Type 1 Diabetes. *Diabetes Technol. Ther.* **2019**, *21*, 286–294. [CrossRef] [PubMed]
4. Adolfsson, P.; Nilsson, S.; Lindblad, B. Continuous glucose monitoring system during physical exercise in adolescents with type 1 diabetes. *Acta Paediatr. Int. J. Paediatr.* **2011**, *100*, 1603–1609. [CrossRef] [PubMed]
5. Herrington, S.J.; Gee, D.L.; Dow, S.D.; Monosky, K.A.; And, E.D.; Pritchett, K.L. Comparison of glucose monitoring methods during steady-state exercise in women. *Nutrients* **2012**, *4*, 1282–1292. [CrossRef] [PubMed]
6. Giani, E.; Macedoni, M.; Barilli, A.; Petitti, A.; Mameli, C.; Bosetti, A.; Cristiano, A.; Radovanovic, D.; Santus, P.; Zuccotti, G.V. Performance of the Flash Glucose Monitoring System during exercise in youth with Type 1 diabetes. *Diabetes Res. Clin. Pract.* **2018**, *146*, 321–329. [CrossRef]

7. Iscoe, K.E.; Campbell, J.E.; Jamnik, V.; Perkins, B.A.; Riddell, M.C. Efficacy of continuous real-time blood glucose monitoring during and after prolonged high-intensity cycling exercise: Spinning with a continuous glucose monitoring system. *Diabetes Technol. Ther.* **2006**, *8*, 627–635. [CrossRef]

8. Laguna, A.J.; Díez, J.L.; Giménez, M.; Bondia, J. Enhanced accuracy of continuous glucose monitoring during exercise through physical activity tracking integration. *Sensors* **2019**, *19*, 3757. [CrossRef]

9. Zaharieva, D.P.; McGaugh, S.; Davis, E.A.; Riddell, M.C. Advances in Exercise, Physical Activity, and Diabetes. *Diabetes Technol. Ther.* **2020**, *22*, S109–S118. [CrossRef]

10. Castle, J.R.; Rodbard, D. How Well Do Continuous Glucose Monitoring Systems Perform during Exercise? *Diabetes Technol. Ther.* **2019**, *21*, 305–309. [CrossRef]

11. Hansen, E.A.; Klee, P.; Dirlewanger, M.; Bouthors, T.; Elowe-Gruau, E.; Stoppa-Vaucher, S.; Phan-Hug, F.; Antoniou, M.C.C.; Pasquier, J.; Dwyer, A.A.; et al. Accuracy, satisfaction and usability of a flash glucose monitoring system among children and adolescents with type 1 diabetes attending a summer camp. *Pediatr. Diabetes* **2018**, *19*, 1276–1284. [CrossRef]

12. Damiano, E.R.; McKeon, K.; El-Khatib, F.H.; Zheng, H.; Nathan, D.M.; Russell, S.J. A comparative effectiveness analysis of three continuous glucose monitors: The Navigator, G4 Platinum, and Enlite. *J. Diabetes Sci. Technol.* **2014**, *8*, 699–708. [CrossRef] [PubMed]

13. National Center for Biotechnology Information. Available online: https://www.ncbi.nlm.nih.gov/ (accessed on 5 October 2020).

14. DerSimonian, R.; Laird, N. Meta-analysis in clinical trials. *Control. Clin. Trials* **1986**, *7*, 177–188. [CrossRef]

15. Hartung, J.; Knapp, G. On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Stat. Med.* **2001**, *20*, 1771–1782. [CrossRef] [PubMed]

16. Lee, C.H.; Cook, S.; Lee, J.S.; Han, B. Comparison of Two Meta-Analysis Methods: Inverse-Variance-Weighted Average and Weighted Sum of Z-Scores. *Genom. Inform.* **2016**, *14*, 173–180. [CrossRef]

17. IntHout, J.; Ioannidis, J.P.; Borm, G.F. The Hartung–Knapp-Sidik–Jonkman method for random effects meta-analysis is straight-forward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med. Res. Methodol.* **2014**, *14*, 25. [CrossRef]

18. van Aert, R.C.; Jackson, D. A new justification of the Hartung–Knapp method for random-effects meta-analysis based on weighted least squares regression. *Res. Synth. Methods* **2019**, *10*, 515–527. [CrossRef]

19. Hedges, L.V. Distribution Theory for Glass's Estimator of Effect Size and Related Estimators Author (s): Larry V. Hedges Published by: American Educational Research Association and American Statistical Association Journaet o Educational Statis6tiL Key Words: Me. *J. Educ. Stat.* **2014**, *6*, 107–128. [CrossRef]

20. Higgins, J.P.; Thompson, S.G.; Spiegelhalter, D.J. A re-evaluation of random-effects meta-analysis. *J. R. Stat. Soc. Ser. A Stat. Soc.* **2009**, *172*, 137–159. [CrossRef]

21. Luo, D.; Wan, X.; Liu, J.; Tong, T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. *Stat. Methods Med. Res.* **2018**, *27*, 1785–1805. [CrossRef]

22. Fayolle, C.; Brun, J.; Bringer, J.; Mercier, J.; Renard, E. Accuracy of continuous subcutaneous glucose monitoring with the GlucoDay® in type 1 diabetic patients treated by subcutaneous insulin infusion during exercise of low versus high intensity. *Diabetes Metab.* **2006**, *32*, 313–320. [CrossRef]

23. Adolfsson, P.; Örnhagen, H.; Jendle, J. The benefits of continuous glucose monitoring and a glucose monitoring schedule in individuals with type 1 diabetes during recreational diving. *J. Diabetes Sci. Technol.* **2008**, *2*, 778–784. [CrossRef] [PubMed]

24. Riddell, M.C.; Milliken, J. Preventing exercise-induced hypoglycaemia in type 1 diabetes using real-time continuous glucose monitoring and a new carbohydrate intake algorithm: An observational field study. *Diabetes Technol. Ther.* **2011**, *13*, 819–825. [CrossRef] [PubMed]

25. Yardley, J.E.; Kenny, G.P.; Perkins, B.A.; Riddell, M.C.; Malcolm, J.; Boulay, P.; Khandwala, F.; Sigal, R.J. Effects of performing resistance exercise before versus after aerobic exercise on glycemia in type 1 diabetes. *Diabetes Care* **2012**, *35*, 669–675. [CrossRef] [PubMed]

26. Kumareswaran, K.; Elleri, D.; Allen, J.M.; Caldwell, K.; Westgate, K.; Brage, S.; Raymond-Barker, P.; Nodale, M.; Wilinska, M.E.; Amiel, S.A.; et al. Physical activity energy expenditure and glucose levels in pregnant women with type 1 diabetes during free-living and controlled conditions. *Diabetes* **2013**, *36*. [CrossRef]

27. Yardley, J.E.; Sigal, R.J.; Kenny, G.P.; Riddell, M.C.; Lovblom, L.E.; Perkins, B.A. Point Accuracy of interstitial continuous glucose monitoring during exercise in type 1 diabetes. *Diabetes Technol. Ther.* **2013**, *15*, 46–49. [CrossRef] [PubMed]

28. Radermecker, R.P.; Fayolle, C.; Brun, J.F.; Bringer, J.; Renard, E. Accuracy assessment of online glucose monitoring by a subcutaneous enzymatic glucose sensor during exercise in patients with type 1 diabetes treated by continuous subcutaneous insulin infusion. *Diabetes Metab.* **2013**, *39*, 258–262. [CrossRef]

29. Yousef, M.; Westman, A.; Lindberg, A.; De Lacerda, C.; Jendle, J. Glucose changes and working memory in individuals with type 1 diabetes during air pressure changes simulating skydiving. *Diabetes Technol. Ther.* **2014**, *16*, 56–62. [CrossRef]

30. Campbell, M.D.; West, D.J.; Bain, S.C.; Kingsley, M.I.; Foley, P.; Kilduff, L.; Turner, D.; Gray, B.; Stephens, J.W.; Bracken, R.M. Simulated games activity vs continuous running exercise: A novel comparison of the glycemic and metabolic responses in T1DM patients. *Scand. J. Med. Sci. Sports* **2015**, *25*, 216–222. [CrossRef]

31. Moser, O.; Tschakert, G.; Mueller, A.; Groeschl, W.; Pieber, T.R.; Obermayer-Pietsch, B.; Koehler, G.; Hofmann, P. Effects of high-intensity interval exercise versus moderate continuous exercise on glucose homoeostasis and hormone response in patients with type 1 diabetes mellitus using novel ultra-long-acting insulin. *PLoS ONE* **2015**, *10*, e0136489. [CrossRef]

32. van Dijk, J.W.; Eijsvogels, T.M.; Nyakayiru, J.; Schreuder, T.H.; Hopman, M.T.; Thijssen, D.H.; van Loon, L.J. Glycemic control during consecutive days with prolonged walking exercise in individuals with type 1 diabetes mellitus. *Diabetes Res. Clin. Pract.* **2016**, *117*, 74–81. [CrossRef]

33. Moser, O.; Mader, J.K.; Tschakert, G.; Mueller, A.; Groeschl, W.; Pieber, T.R.; Koehler, G.; Messerschmidt, J.; Hofmann, P. Accuracy of continuous glucose monitoring (CGM) during continuous and high-intensity interval exercise in patients with type 1 diabetes mellitus. *Nutrients* **2016**, *8*, 489. [CrossRef] [PubMed]

34. Bally, L.; Zueger, T.; Pasi, N.; Carlos, C.; Paganini, D.; Stettler, C. Accuracy of continuous glucose monitoring during differing exercise conditions. *Diabetes Res. Clin. Pract.* **2016**, *112*, 1–5. [CrossRef] [PubMed]

35. Taleb, N.; Emami, A.; Suppere, C.; Messier, V.; Legault, L.; Chiasson, J.L.; Rabasa-Lhoret, R.; Haidar, A. Comparison of Two Continuous Glucose Monitoring Systems, Dexcom G4 Platinum and Medtronic Paradigm Veo Enlite System, at Rest and during Exercise. *Diabetes Technol. Ther.* **2016**, *18*, 561–567. [CrossRef] [PubMed]

36. McAuley, S.A.; Horsburgh, J.C.; Ward, G.M.; La Gerche, A.; Gooley, J.L.; Jenkins, A.J.; MacIsaac, R.J.; O'Neal, D.N. Insulin pump basal adjustment for exercise in type 1 diabetes: A randomised crossover study. *Diabetologia* **2016**, *59*. [CrossRef] [PubMed]

37. Aberer, F.; Hajnsek, M.; Rumpler, M.; Zenz, S.; Baumann, P.M.; Elsayed, H.; Puffing, A.; Treiber, G.; Pieber, T.R.; Sourij, H.; et al. Evaluation of subcutaneous glucose monitoring systems under routine environmental conditions in patients with type 1 diabetes. *Diabetes Obes. Metab.* **2017**, *19*, 1051–1055. [CrossRef]

38. Gawrecki, A.; Naskret, D.; Niedzwiecki, P.; Duda-Sobczak, A.; Araszkiewicz, A.; Zozulinska-Ziolkiewicz, D. High-intensity Exercise in Men with Type 1 Diabetes and Mode of Insulin Therapy. *Int. J. Sports Med.* **2017**, *38*, 329–335. [CrossRef]

39. Jayawardene, D.C.; McAuley, S.A.; Horsburgh, J.C.; Gerche, A.L.; Jenkins, A.J.; Ward, G.M.; MacIsaac, R.J.; Roberts, T.J.; Grosman, B.; Kurtz, N.; et al. Closed-loop insulin delivery for adults with type 1 diabetes undertaking high-intensity interval exercise versus moderate-intensity exercise: A randomized, crossover study. *Diabetes Technol. Ther.* **2017**, *19*, 340–348. [CrossRef]

40. Zaharieva, D.; Yavelberg, L.; Jamnik, V.; Cinar, A.; Turksoy, K.; Riddell, M.C. The effects of basal insulin suspension at the start of exercise on blood glucose levels during continuous versus circuit-based exercise in individuals with type 1 diabetes on continuous subcutaneous insulin infusion. *Diabetes Technol. Ther.* **2017**, *19*, 370–378. [CrossRef]

41. Reddy, R.; El Youssef, J.; Winters-Stone, K.; Branigan, D.; Leitschuh, J.; Castle, J.; Jacobs, P.G.; ElYoussef, J.; Branigan, D.; Castle, J.; et al. The effect of exercise on sleep in adults with type 1 diabetes. *Diabetes Obes. Metab.* **2018**, *20*, 443–447. [CrossRef]

42. Quirós, C.; Bertachi, A.; Giménez, M.; Biagi, L.; Viaplana, J.; Viñals, C.; Vehí, J.; Conget, I.; Bondia, J. Blood glucose monitoring during aerobic and anaerobic physical exercise using a new artificial pancreas system. *Endocrinol. Diabetes Nutr.* **2018**, *65*, 342–347. [CrossRef]

43. Larose, S.; Taleb, N.; Roy-Fleming, A.; Suppere, C.; Messier, V.; Rabasa-Lhoret, R. Comparison of Continuous Glucose Monitoring with Capillary Glucose Levels and Dynamics of Accuracy Changes during Moderate-Intensity Aerobic Exercise in Patients with Type 1 Diabetes. *Can. J. Diabetes* **2018**, *42*, S51–S52. [CrossRef]

44. Aronson, R.; Brown, R.E.; Li, A.; Riddell, M.C. Optimal insulin correction factor in post–high-intensity exercise hyperglycemia in adults with type 1 diabetes: The FIT study. *Diabetes Care* **2019**, *42*, 10–16. [CrossRef] [PubMed]

45. Reddy, R.; Wittenberg, A.; Castle, J.R.; El Youssef, J.; Winters-Stone, K.; Gillingham, M.; Jacobs, P.G. *Effect of Aerobic and Resistance Exercise on Glycemic Control in Adults With Type 1 Diabetes*; Elsevier Inc.: Amsterdam, The Netherlands, 2018; Volume 43, pp. 406–414.e1. [CrossRef]

46. Biagi, L.; Bertachi, A.; Quirós, C.; Giménez, M.; Conget, I.; Bondia, J.; Vehí, J. Accuracy of continuous glucose monitoring before, during, and after aerobic and anaerobic exercise in patients with type 1 diabetes mellitus. *Biosensors* **2018**, *8*, 22. [CrossRef] [PubMed]

47. Abdulrahman, A.; Manhas, J.; Linnane, H.; Gurney, M.; Fitzgerald, C. Annals of Clinical Case Reports The Use of Continuous Glucose Monitoring for Sport in. *Ann. Clin. Case Rep. Diabetol.* **2018**, *3*, 1–5.

48. Castle, J.R.; Youssef, J.E.; Wilson, L.M.; Reddy, R.; Resalat, N.; Branigan, D.; Ramsey, K.; Leitschuh, J.; Rajhbeharrysingh, U.; Senf, B.; et al. Randomized outpatient trial of single- and dual-hormone closed-loop systems that adapt to exercise using wearable sensors. *Diabetes Care* **2018**, *41*. [CrossRef]

49. Moser, O.; Pandis, M.; Aberer, F.; Kojzar, H.; Hochfellner, D.; Elsayed, H.; Motschnig, M.; Augustin, T.; Kreuzer, P.; Pieber, T.R.; et al. A head-to-head comparison of personal and professional continuous glucose monitoring systems in people with type 1 diabetes: Hypoglycaemia remains the weak spot. *Diabetes Obes. Metab.* **2019**, *21*, 1043–1048. [CrossRef]

50. Steineck, I.I.K.; Mahmoudi, Z.; Ranjan, A.; Schmidt, S.; Jørgensen, J.B.; Nørgaard, K. Comparison of Continuous Glucose Monitoring Accuracy between Abdominal and Upper Arm Insertion Sites. *Diabetes Technol. Ther.* **2019**, *21*, 295–302. [CrossRef]

51. Burckhardt, M.A.; Chetty, T.; Smith, G.J.; Adolfsson, P.; De Bock, M.; Jones, T.W.; Davis, E.A. Use of Continuous Glucose Monitoring Trends to Facilitate Exercise in Children with Type 1 Diabetes. *Diabetes Technol. Ther.* **2019**, *21*, 51–55. [CrossRef]

52. Forlenza, G.P.; Buckingham, B.A.; Christiansen, M.P.; Paul Wadwa, R.; Peyser, T.A.; Lee, J.B.; O'Connor, J.; Dassau, E.; Huyett, L.M.; Layne, J.E.; et al. Performance of Omnipod Personalized Model Predictive Control Algorithm with Moderate Intensity Exercise in Adults with Type 1 Diabetes. *Diabetes Technol. Ther.* **2019**, *21*, 265–272. [CrossRef] [PubMed]

53. Larose, S.; Rabasa-Lhoret, R.; Roy-Fleming, A.; Suppère, C.; Tagougui, S.; Messier, V.; Taleb, N. Changes in Accuracy of Continuous Glucose Monitoring Using Dexcom G4 Platinum over the Course of Moderate Intensity Aerobic Exercise in Type 1 Diabetes. *Diabetes Technol. Ther.* **2019**, *21*, 364–369. [CrossRef] [PubMed]

54. Zaharieva, D.P.; Turksoy, K.; McGaugh, S.M.; Pooni, R.; Vienneau, T.; Ly, T.; Riddell, M.C. Lag Time Remains with Newer Real-Time Continuous Glucose Monitoring Technology during Aerobic Exercise in Adults Living with Type 1 Diabetes. *Diabetes Technol. Ther.* **2019**, *21*, 313–321. [CrossRef] [PubMed]

55. Moser, O.; Eckstein, M.L.; Mueller, A.; Birnbaumer, P.; Aberer, F.; Koehler, G.; Sourij, C.; Kojzar, H.; Holler, P.; Simi, H.; et al. Impact of physical exercise on sensor performance of the FreeStyle Libre intermittently viewed continuous glucose monitoring system in people with Type 1 diabetes: A randomized crossover trial. *Diabet. Med.* **2019**, *36*, 606–611. [CrossRef] [PubMed]

56. Zaharieva, D.P.; Cinar, A.; Yavelberg, L.; Jamnik, V.; Riddell, M.C. No Disadvantage to Insulin Pump Off vs Pump On During Intermittent High-Intensity Exercise in Adults With Type 1 Diabetes. *Can. J. Diabetes* **2020**, *44*, 162–168. [CrossRef] [PubMed]

57. Toghi-Eshghi, S.R.; Yardley, J.E. Morning (Fasting) vs Afternoon Resistance Exercise in Individuals with Type 1 Diabetes: A Randomized Crossover Study. *J. Clin. Endocrinol. Metab.* **2019**, *104*, 5217–5224. [CrossRef]

58. Steineck, I.I.; Ranjan, A.; Schmidt, S.; Clausen, T.R.; Holst, J.J.; Nørgaard, K. Preserved glucose response to low-dose glucagon after exercise in insulin-pump-treated individuals with type 1 diabetes: A randomised crossover study. *Diabetologia* **2019**, *62*, 582–592. [CrossRef]

59. Gawrecki, A.; Michalak, A.; Gałczyński, S.; Dachowska, I.; Zozulińska-Ziółkiewicz, D.; Szadkowska, A. Physical workload and glycemia changes during football matches in adolescents with type 1 diabetes can be comparable. *Acta Diabetol.* **2019**, *56*, 1191–1198. [CrossRef]

60. Lee, A.S.; Way, K.L.; Johnson, N.A.; Twigg, S.M. High-intensity interval exercise and hypoglycaemia minimisation in adults with type 1 diabetes: A randomised cross-over trial. *J. Diabetes Complicat.* **2020**, *34*, 107514. [CrossRef]

61. Moser, O.; Eckstein, M.L.; Mueller, A.; Birnbaumer, P.; Aberer, F.; Koehler, G.; Sourij, C.; Kojzar, H.; Pferschy, P.; Dietz, P.; et al. Pre-exercise blood glucose levels determine the amount of orally administered carbohydrates during physical exercise in individuals with type 1 diabetes—A randomized cross-over trial. *Nutrients* **2019**, *11*, 1287. [CrossRef]

62. Scott, S.N.; Cocks, M.; Andrews, R.C.; Narendran, P.; Purewal, T.S.; Cuthbertson, D.J.; Wagenmakers, A.J.; Shepherd, S.O. Fasted high-intensity interval and moderate-intensity exercise do not lead to detrimental 24-hour blood glucose profiles. *J. Clin. Endocrinol. Metab.* **2019**, *104*, 111–117. [CrossRef]

63. Williams, D.; Deere, R.; Moser, O.; McCarthy, O.; Pitt, J.; Bracken, R.; Sourij, H.; Bain, S.; Eckstein, M.; Hayes, J. Performance of the Freestyle Libre flash glucose monitoring (flash GM) system in individuals with type 1 diabetes: A secondary outcome analysis of a randomized crossover trial. *Diabetes Obes. Metab.* **2019**, *21*. [CrossRef]

64. Scott, S.N.; Cocks, M.; Andrews, R.C.; Narendran, P.; Purewal, T.S.; Cuthbertson, D.J.; Wagenmakers, A.J.; Shepherd, S.O. High-Intensity Interval Training Improves Aerobic Capacity Without a Detrimental Decline in Blood Glucose in People with Type 1 Diabetes. *J. Clin. Endocrinol. Metab.* **2018**, *104*, 604–612. [CrossRef] [PubMed]

65. Mccarthy, O.; Eckstein, M.L.; Scott, S.N.; Fontana, F.Y.; Christiansen, M.P.; Stettler, C.; Fisher, M.; Bode, B.; Riddell, M.C.; Hayes, C.; et al. Glycemic responses to strenuous training in male professional cyclists with type 1 diabetes: A prospective observational study. *BMJ Open Diabetes Res. Care* **2020**, *8*, 1–9. [CrossRef] [PubMed]

66. Brockman, N.K.; Sigal, R.J.; Kenny, G.P.; Riddell, M.C.; Perkins, B.A.; Yardley, J.E. Sex-Related Differences in Blood Glucose Responses to Resistance Exercise in Adults With Type 1 Diabetes: A Secondary Data Analysis. *Can. J. Diabetes* **2020**, *44*, 267–273.e1. [CrossRef] [PubMed]

67. Fokkert, M.J.; Dijk, P.R.; Edens, M.A.; Díez, A.; Slingerland, R.J.; Gans, R.O.B.; Delgado, E.; Bilo, H.J.G. Performance of continuous glucose monitoring devices during intensive exercise conditions in people with diabetes: The Mont Blanc experience. *Diabet. Med.* **2020**. [CrossRef] [PubMed]

68. Guillot, F.H.; Jacobs, P.G.; Wilson, L.M.; El Youssef, J.; Gabo, V.B.; Branigan, D.L.; Tyler, N.S.; Ramsey, K.; Riddell, M.C.; Castle, J.R. Accuracy of the dexcom G6 glucose sensor during aerobic, resistance, and interval exercise in adults with type 1 diabetes. *Biosensors* **2020**, *10*, 138. [CrossRef]

69. Mastrototaro, J.; Shin, J.; Marcus, A.; Sulur, G. The accuracy and efficacy of real-time continuous glucose monitoring sensor in patients with type 1 diabetes. *Diabetes Technol. Ther.* **2008**, *10*, 385–390. [CrossRef]

70. Slover, R.H.; Tryggestad, J.B.; DiMeglio, L.A.; Fox, L.A.; Bode, B.W.; Bailey, T.S.; Brazg, R.; Christiansen, M.P.; Sherr, J.L.; Tsalikian, E.; et al. Accuracy of a Fourth-Generation Continuous Glucose Monitoring System in Children and Adolescents with Type 1 Diabetes. *Diabetes Technol. Ther.* **2018**, *20*, 576–584. [CrossRef]

71. Rodbard, D. Continuous Glucose Monitoring: A Review of Successes, Challenges, and Opportunities. *Diabetes Technol. Ther.* **2016**, *18*, S23–S213. [CrossRef]

72. Nakamura, K.; Balo, A. The accuracy and efficacy of the dexcom G4 platinum continuous glucose monitoring system. *J. Diabetes Sci. Technol.* **2015**, *9*, 1021–1026. [CrossRef]

73. Garg, S.K.; Weinzimer, S.A.; Tamborlane, W.V.; Buckingham, B.A.; Bode, B.W.; Bailey, T.S.; Brazg, R.L.; Ilany, J.; Slover, R.H.; Anderson, S.M.; 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*, 155–163. [CrossRef]

74. Gross, T.M.; Bode, B.W.; Einhorn, D.; Kayne, D.M.; Reed, J.H.; White, N.H.; Mastrototaro, J.J. Performance evaluation of the MiniMed® continuous glucose monitoring system during patient home use. *Diabetes Technol. Ther.* **2000**, *2*, 49–56. [CrossRef] [PubMed]

75. Sinha, M.; McKeon, K.M.; Parker, S.; Goergen, L.G.; Zheng, H.; El-Khatib, F.H.; Russell, S.J. A Comparison of Time Delay in Three Continuous Glucose Monitors for Adolescents and Adults. *J. Diabetes Sci. Technol.* **2017**, *11*, 1132–1137. [CrossRef] [PubMed]

76. Higgins JPT, G.S. *Cochrane Handbook for Systematic Reviews of Interventions*; Wiley Blackwell: Hoboken, NJ, USA, 2019.

77. Wan, X.; Wang, W.; Liu, J.; Tong, T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med. Res. Methodol.* **2014**, *14*, 135, [CrossRef] [PubMed]

78. Szadkowska, A.; Gawrecki, A.; Michalak, A.; Zozulinska-Ziokiewicz, D.; Fendler, W.; Młynarski, W. Flash Glucose Measurements in Children with Type 1 Diabetes in Real-Life Settings: To Trust or Not to Trust? *Diabetes Technol. Ther.* **2018**, *20*, 17–24. [CrossRef] [PubMed]

79. Huyett, L.M.; Dassau, E.; Zisser, H.C.; Doyle, F.J. The impact of glucose sensing dynamics on the closed-loop artificial pancreas. In Proceedings of the 2015 American Control Conference (ACC), Chicago, IL, USA, 1–3 July 2015; pp. 5116–5121. [CrossRef]

80. Adolfsson, P.; Örnhagen, H.; Jendle, J. Accuracy and reliability of continuous glucose monitoring in individuals with type 1 diabetes during recreational diving. *Diabetes Technol. Ther.* **2009**, *11*, 493–497. [CrossRef] [PubMed]

81. Riddell, M.C.; Pooni, R.; Yavelberg, L.; Li, Z.; Kollman, C.; Brown, R.E.; Li, A.; Aronson, R. Reproducibility in the cardiometabolic responses to high-intensity interval exercise in adults with type 1 diabetes. *Diabetes Res. Clin. Pract.* **2019**, *148*. [CrossRef]

82. Ekhlaspour, L.; Mondesir, D.; Lautsch, N.; Balliro, C.; Hillard, M.; Magyar, K.; Radocchia, L.G.; Esmaeili, A.; Sinha, M.; Russell, S.J. Comparative Accuracy of 17 Point-of-Care Glucose Meters. *J. Diabetes Sci. Technol.* **2017**, *11*, 558–566. [CrossRef]

83. Danne, T.; Nimri, R.; Battelino, T.; Bergenstal, R.M.; Close, K.L.; DeVries, J.H.; Garg, S.; Heinemann, L.; Hirsch, I.; Amiel, S.A.; et al. International consensus on use of continuous glucose monitoring. *Diabetes Care* **2017**, *40*, 1631–1640. [CrossRef]

84. Reiterer, F.; Polterauer, P.; Schoemaker, M.; Schmelzeisen-Redecker, G.; Freckmann, G.; Heinemann, L.; Del Re, L. Significance and Reliability of MARD for the Accuracy of CGM Systems. *J. Diabetes Sci. Technol.* **2017**, *11*, 59–67. [CrossRef]