Meta-analysis

Effectiveness of artificial pancreas in the non-adult population: A systematic review and network meta-analysis

Vasilios Karageorgiou, Theodoros G. Papaioannou, Ioannis Bellos, Krystallenia Alexandraki, Nikolaos Tentolouris, Christodoulos Stefanidis, George P. Chrouso, Dimitrios Tousoulis

Abstract

Objective: Artificial pancreas is a technology that minimizes user input by bridging continuous glucose monitoring and insulin pump treatment, and has proven safety in the adult population. The purpose of this systematic review and meta-analysis is to evaluate the efficacy of closed-loop (CL) systems in the glycemic control of non-adult type 1 diabetes patients in both a pairwise and network meta-analysis (NMA) context and investigate various parameters potentially affecting the outcome.

Methods: Literature was systematically searched using the MEDLINE (1966–2018), Scopus (2004–2018), Cochrane Central Register of Controlled Trials (CENTRAL) (1999–2018), Clinicaltrials.gov (2008–2018) and Google Scholar databases. Studies comparing the glycemic control in CL (either single- or dual-hormone) with continuous subcutaneous insulin infusion (CSII) in people with diabetes (PWD) aged <18 years old were deemed eligible. The primary outcome analysis was conducted with regard to time spent in the target glycemic range. All outcomes were evaluated in NMA in order to investigate potential between-algorithm differences. Pairwise meta-analysis and meta-regression were performed using the RevMan 5.3 and Open Meta-Analyst software. For NMA, the package pcnetmetain R 3.5.1 was used.

Results: The meta-analysis was based on 25 studies with a total of 504 PWD. The CL group was associated with significantly higher percentage of time spent in the target glycemic range (Mean (SD): 67.59% (SD: 8.07%) in the target range and OL PWD spending 55.77% (SD: 11.73%), MD: −11.97%, 95% CI [−18.40, −5.54%]) and with lower percentages of time in hyperglycemia (MD: 3.01%, 95% CI [1.68, 4.34%]) and hypoglycemia (MD: 0.67%, 95% CI [0.21, 1.13%]). Mean glucose was also decreased in the CL group (MD: 0.75 mmol/L, 95% CI [0.18–1.33]). The NMA arm of the study showed that the bihormonal modality was superior to other algorithms and standard treatment in lowering mean glucose and increasing time spent in the target range. The DiAs platform was superior to PID in controlling hypoglycemia and mean glucose. Time in target range and mean glucose were unaffected by the confounding factors tested.

Conclusions: The findings of this meta-analysis suggest that artificial pancreas systems are superior to the standard sensor-augmented pump treatment of type 1 diabetes mellitus in non-adult PWD. Between-algorithm differences are also addressed, implying a superiority of the bihormonal treatment modality. Future large-scale studies are needed in the field to verify these outcomes and to determine the optimal algorithm to be used in the clinical setting.

Keywords: Artificial pancreas
Closed-loop
Diabetes mellitus
Pediatric
Meta-analysis

ARTICLE INFO

Article history:
Received 31 July 2018
Accepted 9 October 2018

Keywords:
Artificial pancreas
Closed-loop
Diabetes mellitus
Pediatric
Meta-analysis

ABSTRACT

Objective: Artificial pancreas is a technology that minimizes user input by bridging continuous glucose monitoring and insulin pump treatment, and has proven safety in the adult population. The purpose of this systematic review and meta-analysis is to evaluate the efficacy of closed-loop (CL) systems in the glycemic control of non-adult type 1 diabetes patients in both a pairwise and network meta-analysis (NMA) context and investigate various parameters potentially affecting the outcome.

Methods: Literature was systematically searched using the MEDLINE (1966–2018), Scopus (2004–2018), Cochrane Central Register of Controlled Trials (CENTRAL) (1999–2018), Clinicaltrials.gov (2008–2018) and Google Scholar (2004–2018) databases. Studies comparing the glycemic control in CL (either single- or dual-hormone) with continuous subcutaneous insulin infusion (CSII) in people with diabetes (PWD) aged <18 years old were deemed eligible. The primary outcome analysis was conducted with regard to time spent in the target glycemic range. All outcomes were evaluated in NMA in order to investigate potential between-algorithm differences. Pairwise meta-analysis and meta-regression were performed using the RevMan 5.3 and Open Meta-Analyst software. For NMA, the package pcnetmetain R 3.5.1 was used.

Results: The meta-analysis was based on 25 studies with a total of 504 PWD. The CL group was associated with significantly higher percentage of time spent in the target glycemic range (Mean (SD): 67.59% (SD: 8.07%) in the target range and OL PWD spending 55.77% (SD: 11.73%), MD: −11.97%, 95% CI [−18.40, −5.54%]) and with lower percentages of time in hyperglycemia (MD: 3.01%, 95% CI [1.68, 4.34%]) and hypoglycemia (MD: 0.67%, 95% CI [0.21, 1.13%]). Mean glucose was also decreased in the CL group (MD: 0.75 mmol/L, 95% CI [0.18–1.33]). The NMA arm of the study showed that the bihormonal modality was superior to other algorithms and standard treatment in lowering mean glucose and increasing time spent in the target range. The DiAs platform was superior to PID in controlling hypoglycemia and mean glucose. Time in target range and mean glucose were unaffected by the confounding factors tested.

Conclusions: The findings of this meta-analysis suggest that artificial pancreas systems are superior to the standard sensor-augmented pump treatment of type 1 diabetes mellitus in non-adult PWD. Between-algorithm differences are also addressed, implying a superiority of the bihormonal treatment modality. Future large-scale studies are needed in the field to verify these outcomes and to determine the optimal algorithm to be used in the clinical setting.

© 2018 Elsevier Inc. All rights reserved.

Contents

1. Introduction ............................................................... 21
2. Materials and Methods .......................................................... 21
   2.1. Study Design ............................................................ 21
   2.2. Literature Search and Data Collection ................................. 21

* Corresponding author at: Biomedical Engineering Unit, 1st University Department of Cardiology, Hippokration General Hospital of Athens, 114 Vas. Sophias ave, Athens 11527, Greece. E-mail address: thepap@med.uoa.gr (T.G. Papaioannou).

0026-0495/© 2018 Elsevier Inc. All rights reserved.
1. Introduction

Type 1 diabetes mellitus (T1D) accounts for 5–10% of cases of diabetes worldwide, whose US prevalence has increased to 1.93 per 1000 during the last decade [1]. The diagnosis in people with diabetes (PWD) is made primarily during their childhood and adolescence [2]. It almost invariably results in exogenous insulin administration for proper management [3], a treatment modality demanding significant therapeutic input through CGM [7]. Bihormonal algorithm control relies on both insulin and glucagon infusion [8]. A number of clinical trials have reported the safety and efficacy in minimizing hypoglycemia and hyperglycemia in children and adolescent outpatients [4].

The technology of the artificial pancreas (AP) employs an algorithm (of various structures) that bridges the CGM device with the insulin pump, thereby independently (without input from the user) determining the dose of the insulin needed. Four algorithms used in AP are proportional-integral-derivative (PID) algorithms [5], model predictive control (MPC) algorithms [6], fuzzy logic algorithms [7] and bihormonal algorithms [8]. Schematically, PID algorithms measure glucose and modify insulin infusion rate according to the sampled value’s difference from the glucose target point as expressed by the percentage of time spent in the glycemic range, mean glucose) and demographic parameters (country, language or date restrictions were applied. The date of the last search was set at 20 April 2018. The eligibility criteria included pediatric patients (of various ages) and studies comparing the glycemic control of CL (either single- or dual-hormone) with continuous subcutaneous insulin infusion (CSII) in PWD aged <18 years old were deemed eligible. The choice of a certain algorithm over another is not supported by direct, i.e. clinical trial, evidence so NMA is an appropriate choice for suggesting the possible superiority of an algorithm group. To our knowledge, this is the first time that a network meta-analysis is done exclusively on the non-adult population, while possible sources of heterogeneity are investigated by examining various demographic and intervention characteristics as potential confounders in a meta-regression analysis context.

2. Materials and Methods

2.1. Study Design

The current systematic review and meta-analysis are in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) [14]. All studies comparing the glycemic control of CL (either single- or dual-hormone) with continuous subcutaneous insulin infusion (CSII) in PWD aged <18 years old were deemed eligible. The studies were selected in three consecutive stages. Firstly, the titles and/or abstracts of all electronic articles were screened and subsequently articles that were presumed to meet the criteria were retrieved as full texts. Finally, all studies reporting the outcome of interest were included in this systematic review. Review articles, animal studies, in silico simulations, safety studies, non-comparative studies (case reports, case series) and studies in patients aged >18 years old were excluded. Any discrepancies regarding the methodology, retrieval of articles and statistical analysis were resolved through the consensus of all authors.

2.2. Literature Search and Data Collection

The literature search was conducted based on the algorithm: (artificial pancreas OR closed loop) AND (adolescents OR children OR kids OR youth), which was applied in the MEDLINE (1966–2018), Scopus (2004–2018), Cochrane Central Register of Controlled Trials (CENTRAL) (1999–2018) and ClinicalTrials.gov (2008–2018) databases and Google Scholar (2004–2018). Also, the reference lists of the included studies were screened (snow-ball method), to identify additional sources. No language or date restrictions were applied. The date of the last search was set at 20 April 2018. The flowchart of the literature search is presented in Fig. 1. The extracted data from each study included the following: name of the first author, date of publication, primary and secondary endpoints (time in target range, time in hypoglycemia, time in hyperglycemia, mean glucose) and demographic parameters (country,
age, sex, body mass index (BMI), glycated hemoglobin A1c (HbA1c) at presentation, duration of diabetes, insulin pump use).

2.3. Quality Assessment

The methodologic quality of each clinical trial was evaluated using both the Jadad scale [15] and the Cochrane risk of bias tool for visualization purposes. Two researchers (V.K., I.B.) independently evaluated the studies on the grounds of possible selection bias (random sequence generation), detection bias (blinding of outcome assessment), performance bias (blinding of participants and personnel), attrition bias (incomplete outcome data) and reporting bias (selective reporting), as well as an overall assessment of the risk of bias (other bias). Potential disagreements were resolved by the consensus of all authors.

2.4. Statistical Analysis

Statistical meta-analysis was performed with the RevMan 5.3 software (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011). Confidence intervals were set at 95%. The inter-study heterogeneity was assessed using the inconsistency index ($I^2$) [16]. When significant heterogeneity was present ($I^2 > 50$%), the DerSimonian–Laird random effect model was applied to provide pooled estimates of the mean difference (MD) and the 95% confidence intervals (95% CI). For the outcomes discussed, unweighted means and standard deviations (SD) of the included studies are reported for reference. The primary quantitative synthesis was planned to estimate the mean difference of percentage of time spent in target range (4.0–10 mmol/L) in a CL setting versus time in CSII. Subgroup analysis was conducted on the basis of meal or exercise announcement. Secondary analyses included the comparisons of mean glucose, as well as of percentages of time spent in hypoglycemia ($<4$ mmol/L) and hyperglycemia. Publication bias was evaluated by the visual inspection of funnel plots, since the high inter-study heterogeneity precluded the safe interpretation of the available statistical tests [17]. Moreover, to determine potential sources of heterogeneity, univariate meta-regression was performed, using the Open Meta- Analyst statistical software [18]. Specifically, the effects of HbA1c, age, gender, body mass index, diabetes duration, time of intervention, total daily insulin, type of algorithm, outpatient setting, meal/exercise announcement, study randomization and year of publication were explored.

Fig. 1. Flow diagram depicting the information process across the stages of the systematic review.
For the NMA, the package pcnetmeta in R 3.5.1 was used [19]. Arm-based analysis was performed using Markov Chain Monte Carlo Convergence (MCMC) simulation. The number of iterations for the adaptation process was set at the default 5000. For the demonstration of between-algorithm differences, posterior density plot curves were generated. Considering time in hyperglycemia (Suppl. Table 1), the differing definition of hyperglycemia would give rise to non-realistic results. Data that express duration above differing concentrations cannot be considered comparable for such an analysis. Thus, similarity among studies cannot be assumed and so a NMA for time in hyperglycemia was not performed [20].

2.5. Sensitivity Analysis

The influence of individual studies was explored by performing leave-one-out analyses; one study was sequentially omitted at a time in order to find out its effect in the overall outcome. The software used to conduct this analysis was Open Meta-Analyst. As significant heterogeneity was present, a graphical display of study heterogeneity (GOSH) plot was constructed. This exploratory graphical representation plots the results of every possible subset of included studies against $I^2$ to detect possible sources of heterogeneity as clusters that disrupt the normal distribution [21]. Identification of each individual study’s effect on heterogeneity was done with Baujat plots. This tool plots the effect of each individual study on Cochran Q-test for heterogeneity against the standardized square difference of the treatment effect, in our case, mean difference, as calculated with and without the study’s data [22]. For these plots, the package metafor in R 3.5.1 was used [23]. After studies have been identified and excluded, a repeat GOSH plot and meta-analysis will be performed in order to assess the heterogeneity and the mean difference.

3. Results

3.1. Included Studies

Twenty-five studies [9–12,24–44] were finally included in this review, with a total of 504 patients. The methodologic characteristics of the included studies (country, exclusion criteria, study design, time of intervention, algorithm, definition of hyperglycemia, inpatient/outpatient setting, meal/exercise announcement) are described in Suppl. Table 1. Suppl. Table 2 presents the main PWD characteristics (number, age, gender, weight, body mass index, T1D duration, HbA1c and total daily insulin). Quantitative synthesis consisted of 19 studies [6,7,14–19,21–25,27–34] while six studies [9,10,12,26,30,36] were only included in the qualitative synthesis, since all their outcomes were reported in terms of median and interquartile range.

3.2. Excluded Studies

Six studies [45–50] were excluded after reading the full-text. Five of them lacked an open-loop control group, but conducted their comparisons between two closed-loop groups, investigating the effects of predictive hyperglycemia and hypoglycemia minimization system [45,46], pramlintide administration [47], heart-rate triggered algorithm calibration [48] and snacking [49], respectively. Also, one study [50] was not included because it did not provide the standard deviation for the percentage time in target range.

3.3. Quality Assessment

The outcomes of the Jadad score are presented in Suppl. Table 1. Eight studies (32%) scored 3 points, 8 studies (32%) 2 points, 3 studies (12%) 1 point and 6 studies (24%) 0 points. As depicted in Fig. 2, the lack of a double-blinded design may have led to high risk of bias in the domains of allocation concealment, blinding of participants and personnel and blinding of outcome assessment.

3.4. Qualitative Synthesis

The outcomes of the qualitative synthesis are shown in Suppl. Table 3. It is evident that in all studies, time in target range was significantly superior in the CL group. This finding is less uniform regarding time spent in hypoglycemia, since only three [5,6,30] out of seven comparisons yielded a significant reduction for the CL group. Time in hyperglycemia [5,14,16,21,26] and mean glucose [5,14,16,21,26] were significantly reduced for the CL group in 6 comparisons in a total of 9 and 8 comparisons, respectively.

3.5. Quantitative Synthesis

3.5.1. Pairwise Meta-analyses

The outcome of the pairwise meta-analysis is illustrated in Fig. 3. Specifically, the closed loop group exhibited a significantly increased percentage of time in the target glycemic range (MD: $-11.97\%$, 95% CI $[-18.40, -5.54\%]$, 18 comparisons), with CL PWD spending a mean of 67.59% (SD: 8.07) in the target range and OL PWD spending 55.77% (SD: 11.73). Subgroup analysis showed that this difference was also significant in both groups of announced (Mean SD): 57.55% (8.11) for OL, 68.73% (6.02) for CL MD: $-11.29\%$, 95% CI $[-18.65, -3.92\%]$) and unannounced meal or exercise (Mean: 46.87% (20.02) for OL, 61.9% (13.08) for CL, MD: $-15.36\%$, 95% CI $[-27.83, -2.89\%]$). The results of the secondary analysis indicated that the open-loop group was associated with significantly higher mean glucose values (Means SD): 9.3 mmol (1.32) for OL, 8.59 mmol/L (0.84) for CL, MD: 0.75 mmol/L (95% CI [0.18, 1.33 mmol/L], 20 comparisons), as well as higher percentage of time in hypoglycemia (Means SD): 2.75% (1.43) for OL, 1.71% (1.13) for CL; MD: 0.67%, 95% CI [0.21, 1.13], 11 comparisons) and hyperglycemia (Means SD): 14.07% (13.6) for OL, 7.54% (6.6) for CL, MD: 3.01%, 95% CI [1.68, 4.34%], 11 comparisons) (Fig. 4). The visual inspection of funnel plots revealed asymmetry and thus the possibility of publication bias cannot be excluded (Suppl. Fig. 1).

The univariate meta-regression analysis showed that the covariates examined (HbA1c, diabetes duration, time of intervention, total daily insulin, age, gender, BMI, year of publication, randomization, algorithm type, outpatient setting and meal announcement) did not exert significant effect on the primary outcome of the present meta-analysis (Table 1). Considering the confounding factors affecting time in hypoglycemia, it is shown that there is a significant influence of the duration of the intervention, year of publication, algorithm type and presence of unannounced meals or exercise on the mean difference (Suppl. Table 4). Specifically, studies that featured a closed-loop arm of $<72$ h exhibited a significantly lower mean difference of percentage time in hypoglycemia compared with studies of $>72$ h. In addition, studies utilizing a PID algorithm design exhibited a significantly lower mean difference. The PID-based studies were also inferior concerning the percentage time in hyperglycemia (Suppl. Table 5). Similarly, studies featuring an unannounced meal or exercise also exhibited an inferior control of hyperglycemia compared with the control group, but still the closed loop group spent significantly less time in hyperglycemia. The only factor that affected mean difference in mean glucose was the presence of an unannounced meal or exercise (Suppl. Table 6).

3.6. Network Meta-analysis

Eighteen trials were included in the NMA. Three analyses for time in target range (primary outcome), time in hypoglycemia and mean glucose were conducted. The primary outcome analysis involved 5 different treatment modalities (SAP, MPC, PID, Bihormonal, DiAs). The results of this analysis are depicted in Fig. 5. Of the 10 possible unique study designs (e.g. SAP vs. bihormonal, MPC vs. PID), only 4 (40%) were studied in a direct head-to-head fashion. All trials (100%) used an SAP control group, 2 (11.11%, $n = 51$ PWD) used a bihormonal arm, 5 used a DiAs arm (27.78%, $n = 123$ PWD), 6 used an MPC arm.
(33.33%, n = 101) and 5 used a PID arm (27.78%, n = 98 PWD). PWD data for MD-Logic AP performance was eligible only for the mean glucose analysis.

Table 2 demonstrates the results of the analysis. All algorithm groups demonstrate a larger effect size in percentage time in target range compared with SAP (Table 2). In the posterior density plot (Fig. 6a), which is a visualization of varying treatment effects, it is shown that the bihormonal treatment group has no overlap range with SAP. The numerical results indicate that bihormonal, PID, DiAs and MPC are all, in descending order according to effect size, more effective than SAP treatment. The indirect evidence from comparisons between algorithms, e.g. MPC vs. bihormonal, as depicted in (Table 2) shows that the bihormonal modality may be significantly more efficacious in keeping a target glucose concentration compared with DiAs, MPC and PID.

The NMA for percentage time in hypoglycemia showed that all algorithm groups were more efficient in reducing time in hypoglycemia (Table 2). Only the DiAs modality showed satisfactory confidence intervals (−1.41%, 95% CI: −1.97, −0.73). Indirect evidence also suggests that the DiAs modality is significantly more effective compared with PID (MD: −1.12%, 95% CI: −1.97, −0.34). In the posterior density plot (Fig. 6b), the bihormonal modality appears to have a small overlap region with SAP but the scarce patient data (n = 51 PWD), which is visible as low height of the curve in the posterior density plot, precludes the drawal of a safe conclusion.

In the mean glucose NMA, all algorithm groups showed a reduction in mean glucose compared with SAP, with bihormonal and DiAs modalities reaching significance (−2.17 mmol/L, 95% CI: −2.68, −1.30 and −0.80 mmol/L, 95% CI: −1.18, −0.43 respectively). Through the indirect comparisons generated, the bihormonal algorithm showed a significant superiority compared with DiAs, MPC and PID algorithms (BIH vs. DiAs: −1.38 mmol/L, 95% CI: −2.07, −0.63, BIH vs. MPC: −1.48%, 95% CI: −2.39, −0.46, BIH vs. PID: −1.75 mmol/L, 95% CI: −2.34, −1.16). In addition, the DiAs platform showed marginal superiority compared with PID.

3.7. Sensitivity Analysis

The outcomes of the leave-one-out analysis are depicted in Suppl. Fig. 2. The results for time in target and mean glucose remained stable, as they were not significantly altered by the exclusion of any study. Regarding time in hypoglycemia, omission of one study [29] lowered the mean difference (MD: 0.44%, 95% CI: [0.04, 0.84%]), as well as the heterogeneity ($I^2 = 57$%). In the analysis of time in hyperglycemia, the estimated pooled effect size was moderately affected by one study [41], but still the mean difference in percentage time in hyperglycemia remained statistically significant.

Due to significant heterogeneity ($I^2 = 96%$) in our primary outcome analysis, a GOSH plot was constructed [21], plotting $2^{18} - 1 = 262,143$ possible combinations of subsets. The graph (Suppl. Fig. 3a) shows a bimodal distribution and a mean $I^2$ of 93%, i.e. significant heterogeneity caused by a population. In this case, as suggested by Olkin et al., the studies that influence the result should be dropped out and a repeat analysis should be performed. A Baujat plot was constructed in order to investigate the effect of each individual study on Cochran Q-test for heterogeneity. The result is shown in Suppl. Fig. 4. After excluding 7 studies [28,29,32,37,38,40,41] of the top right quartile, we could reach a homogeneous result ($I^2 = 0$%). In this analysis of 11 studies, CL PWD spent 11.85% more time in the target range [95% CI: 9.57, 14.13] (Suppl. Fig. 5). A repeat GOSH plot was generated and the results show unimodality in both axes and a mean $I^2 = 3.36$, which supports homogeneity (Suppl. Fig. 3b).

The same strategy was followed for all the outcomes. In time in hypoglycemia analysis, the exclusion of 3 studies [29,35,40] showed a more homogeneous ($I^2 = 28$%) but still significant mean difference favoring CL (Means (SD): 2.79% (1.43) for OL, 1.71% (1.13) for CL; MD: 0.76%, 95% CI: 0.25, 1.27). For time in hyperglycemia, the exclusion of
4 studies [32,38,41,43] also resulted in a significantly lower time in hyperglycemia for the CL subjects (Means (SD): 11.1% (7.43) for OL, 6.24% (4.07) for CL, MD: 2.96%, 95% CI: 0.76, 5.17, $I^2 = 40$%). Mean glucose also remained significantly lower in the CL subjects (Means (SD): 7.9% (5.75) for OL, 6.3% (4.17) for CL, MD: 1.6%, 95% CI: 0.42, 2.78, $I^2 = 40$%). The reduction of mean glucose during the CL arm of the studies ($−0.75$ mmol/L, 95% CI: $−1.33$, $−0.18$ mmol/L) remained unaffected by all the potential confounding factors tested, except the unannounced meal or exercise design. As far as time in hypoglycemia is concerned, the outcome indicated a more modest reduction ($−1.33$, 95% CI: $−2.33$, $−0.33$) than previously described ($−1.58$, 95% CI: $−3.66$, $−0.50$) [13]. The meta-regression analysis challenges the robustness of this finding, since significant influence of various parameters (time of intervention, gender, year of publication, unannounced meal or exercise) was noted. Specifically, studies in which the artificial pancreas arm lasted $<72$ h indicated a stronger superiority of the artificial pancreas. In addition, the qualitative synthesis further supports the cautious interpretation of the reduction of time in hypoglycemia, as only 3 out of 9 comparisons yielded a significant benefit for the CL group (Suppl. Table 3).

On the other hand, the marked heterogeneity of the time in hypoglycemia outcome ($I^2 = 95$%) can be partially attributed to the differing definition of hypoglycemia among the studies included, as presented in Suppl. Table 1. By excluding the studies that were most influential on heterogeneity, a repeat analysis also indicated superiority of CL in reducing time in hyperglycemia. As proposed by Bekiar et al. [54], a common repository of data on future trials might resolve this issue.

In the NMA part of our study, the superiority of algorithm treatment modalities over SAP is shown in a) percentage time spent in the target glycemic range, b) percentage time spent in hypoglycemia and c) mean glucose. Statistical significance was not reached.
Fig. 4. Forest plots for secondary outcomes. A) Mean glucose, B) percentage time spent in hypoglycemia (<4 mmol/L), C) percentage time spent in hyperglycemia. CL: closed-loop; OL: open loop.

Table 1
Meta-regression analysis for the percentage time spent in target glycemic range. MPC: model predictive control; PID: proportional integral derivative; DiAs: diabetes assistant; BMI: body mass index; TDI: total daily insulin.

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Level</th>
<th>Number of studies</th>
<th>Coefficient</th>
<th>Standard deviation</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1C</td>
<td>–</td>
<td>18</td>
<td>-2.640</td>
<td>5.704</td>
<td>0.644</td>
</tr>
<tr>
<td>Diabetes duration</td>
<td>–</td>
<td>17</td>
<td>1.572</td>
<td>1.205</td>
<td>0.192</td>
</tr>
<tr>
<td>Time of intervention</td>
<td>–</td>
<td>18</td>
<td>0.001</td>
<td>0.005</td>
<td>0.885</td>
</tr>
<tr>
<td>TDI</td>
<td>–</td>
<td>13</td>
<td>42.105</td>
<td>41.468</td>
<td>0.307</td>
</tr>
<tr>
<td>Age</td>
<td>–</td>
<td>17</td>
<td>0.668</td>
<td>0.617</td>
<td>0.279</td>
</tr>
<tr>
<td>Gender</td>
<td>–</td>
<td>16</td>
<td>-0.146</td>
<td>0.266</td>
<td>0.582</td>
</tr>
<tr>
<td>BMI</td>
<td>–</td>
<td>7</td>
<td>-1.645</td>
<td>1.442</td>
<td>0.254</td>
</tr>
<tr>
<td>Year of publication</td>
<td>–</td>
<td>18</td>
<td>0.835</td>
<td>1.452</td>
<td>0.365</td>
</tr>
<tr>
<td>Algorithm</td>
<td>Modulara</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>MPC</td>
<td>6</td>
<td>4.361</td>
<td>8.612</td>
<td>0.613</td>
</tr>
<tr>
<td></td>
<td>PID</td>
<td>5</td>
<td>5.276</td>
<td>8.787</td>
<td>0.548</td>
</tr>
<tr>
<td></td>
<td>DiAs</td>
<td>4</td>
<td>-6.463</td>
<td>8.908</td>
<td>0.468</td>
</tr>
<tr>
<td></td>
<td>Bihormonal</td>
<td>2</td>
<td>-4.084</td>
<td>9.721</td>
<td>0.674</td>
</tr>
<tr>
<td>Setting</td>
<td>Outpatienta</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Inpatient</td>
<td>13</td>
<td>6.691</td>
<td>4.962</td>
<td>0.178</td>
</tr>
<tr>
<td>Unannounced meal/exercise</td>
<td>Noa</td>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Randomized</td>
<td>Yes</td>
<td>3</td>
<td>-4.075</td>
<td>6.229</td>
<td>0.513</td>
</tr>
<tr>
<td></td>
<td>Yesa</td>
<td>14</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>4</td>
<td>6.922</td>
<td>5.444</td>
<td>0.203</td>
</tr>
</tbody>
</table>

a Reference variable.
for MPC (time in target, hypoglycemia, mean glucose). Still, the density curve indicates a therapeutic effect that warrants further investigation (Fig. 6).

In the indirect comparison arm, it is noteworthy that the bionormal algorithm was significantly superior to all other groups in the increase of time spent in target glycemic range and mean glucose reduction. Other results that indirectly provide an informed estimation of head-to-head comparisons between algorithms are the superiority of DiAs platform over PID in reducing time spent in hypoglycemia (−1.12%, 95% CI: −1.97, −0.34) and mean glucose (−1.11 mmol/L, 95% CI: −1.53, −0.51).

4.1. Strengths and Limitations of the Study

Given the already established safety, feasibility and superiority of the artificial pancreas system in glycemic control, as shown in three systematic reviews and meta-analyses [13, 54, 55], our study sought to recreate and clarify this effect in the pediatric population, which appeared problematic with regards to its heterogeneity [13]. Considering the strengths of our study, the search strategy was broad and encompassed a significant volume of the literature, while the exclusion criteria were strict. Compared to the other systematic reviews on artificial pancreas technology [55], our study included the largest sample of pediatric patients (n = 504). Furthermore, it should be highlighted that, considering time in target range, the superiority of CL control was unaffected by differences in age, sex, BMI, duration of diabetes at the time of intervention, total daily insulin dosage, year of publication, meal announcement, algorithm and outpatient or inpatient study design. As a result, the risk of confounding is limited and therefore the outcomes can be characterized as robust. We also performed a NMA assuming each algorithm as a distinct treatment modality, which allowed us to perform indirect comparisons between algorithms. To our knowledge, this is the first time that between-algorithm differences are systematically treated in a NMA context, suggesting significant differences, and demographic parameters and study characteristics are treated as potential confounders in the artificial pancreas glycemic control efficacy.

The limitations of our study include the relatively small number of PWD included and the heterogeneity of the primary and secondary outcomes, which is only partially explained by the unannounced meal or exercise parameters. We tried to address the issue of heterogeneity by identifying the studies that most affected the result, excluding them and repeating the analysis. Results remained significant and favored the CL modality. Three studies [28, 32, 37] did not favor the CL intervention regarding percent time spent in hyperglycemia. However, in two of these studies [28, 32], a crossover design was not adopted and the two study groups may have received a different level of care. Del Favero et al. [37] explain the apparent inferiority of the CL (p = 0.022) as a result of a) lower HbA1c in the OL arm of the study at the cost of increased hypoglycemic events, and b) cautious algorithm adjustment considering the hypoglycemia correction. Sharifi et al. [28] attributed the result to an altered behavior of PWD and their parents with regards to meal or exercise choices, presumably due to the study process per se. NMA could not be conducted for hyperglycemia due to differing reporting of the outcome and thus similarity for this outcome between studies could not be presupposed [20]. Limited information was also available regarding the HbA1c status post-intervention, due to the nature of the HbA1c kinetics. This glycosylated molecule reflects the mean glucose exposure in the 4 weeks before the analysis rather than that of the last week [56]. The only available information comes from one study [25], which reported a non-significant decrease (−0.3%, 95% CI: [−0.6, 0.1], p-value = 0.17) in a 12-week intervention interval in 25 participants. Moreover, it is noteworthy that no study recruited PWD having a diagnosis of T1D for less than a year or patients that...
were not using insulin pump or used it for <3 months. Therefore, the above results cannot be safely generalized to this dynamic population that exhibits differing insulin needs (minimum 3–6 months after diagnosis) [57]. In this population, CGM alone does not have a clear benefit over standard self-monitoring and insulin therapy [4]. Our study also focused on specific clinical outcomes of the CL treatment modalities as reflected in clinical trials in the non-adult population. Other experimental, technological or in silico advances were not analyzed in depth. In addition, the impossibility of a double-blind design in a CL vs. OL setting constitutes an inherent limitation of all studies evaluating artificial pancreas efficacy.

4.2. Implications for Current Clinical Practice and Future Research

This systematic review suggests the superiority of artificial pancreas systems in the glycemic control of the pediatric population; however, several aspects remain to be elucidated before this technology can be applied widely in clinical practice. Large-scale clinical trials with longer follow-up periods are needed to fully clarify the glycemic benefit that comes from the implementation of closed-loop systems and to investigate its effect on glucose variability, hypoglycemia risk, HbA1c levels, as well as on acute and chronic diabetes complications. Moreover, as revealed in this study, possible confounding factors, such as time of intervention and meal announcement, should be taken into consideration, while head-to-head comparisons of different algorithms and clinical settings should be conducted. These head-to-head trials would clarify our indirect comparison findings that suggest a superiority of the bihormonal modality, among other between-algorithm differences in efficacy. Additionally, broader patient populations, such as PWD with T1D diagnosed for less than a year and PWD with end-organ damage should be included, to further elucidate the extent of CL system efficacy. Finally, parameters as quality of life and need for hospitalization should be evaluated, whereas economic analyses are necessary to assess the cost-effectiveness of closed-loop systems in the management of T1D pediatric PWD.

5. Conclusion

This meta-analysis suggests that the artificial pancreas systems are superior to the standard sensor-augmented pump treatment of type 1 diabetes mellitus in non-adult PWD, significantly increasing the percentage of time spent in the target range and decreasing time in hypoglycemia and hyperglycemia. Before closed-loop systems can be widely implemented in the clinical setting, future large-scale trials should verify these outcomes and provide data regarding the effect of different algorithms, as well as the cost-effectiveness of artificial pancreas. These studies should take into account the influence of potential confounders that were mentioned in this meta-analysis, in order to limit heterogeneity and give a realistic estimation of artificial pancreas efficacy.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.metabol.2018.10.002.

Conflict of Interest

All authors have no conflicts of interest to disclose.

Funding Source

No funding was secured for this study.


