PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Effectiveness, safety and costs of the FreeStyle Libre glucose monitoring system for children and adolescents with type 1 diabetes in Spain: a prospective, uncontrolled, pre-post study
AUTHORS	González-Pacheco, Himar; Rivero-Santana, Amado; Ramallo Farina, Yolanda; Valcárcel-Nazco, Cristina; Álvarez-Pérez, Yolanda; García-Pérez, Lidia; García-Bello, Miguel Angel; Perestelo-Pérez, Lilisbeth; Serrano-Aguilar, Pedro; Group, Health Professional

VERSION 1 – REVIEW

REVIEWER	Michaud, Tzeyu University of Nebraska Medical Center, Department of Health Promotion
REVIEW RETURNED	03-Apr-2023

GENERAL COMMENTS	The manuscript sought to examine the effectiveness, safety and costs of FSL glucose monitoring system among children and adolescent with T1M. While this study may have some merits, the study design is not clear, lacking of sufficient details in a structured/clear format. For example, it is clear either in the Abstract or in the Methods section whether authors recruited participants who were already using FSL (thus the observation study) or recruited them to participate in the FSL intervention (thus a pre-post study design). In addition, there are significant amounts of grammar issues. Authors may benefit from the assistance of scientific writing service. Below are some comments (but not all), as it is not worth the effort to review the Results and Discussion sections without addressing the aforementioned issues first. Abstract. The abstract section should provide sufficient details for readers with an overview of the study without further looking into information in the main text. Apparently, authors have failed to do so. Some examples include, but not limited to: 1. The reason why the number of n=165 recruited participants was reduced to n=156 participants included in the analysis was not clear. 2. Need more details in terms of how costs were measured (instead of just the data source) and what are included. 3. Authors should align the outcomes reported in the Results section with the Outcome measures section. Several outcomes were not specified their operationalization, such as server hypoglycemic events, sensor usage time and scan, device adherence. Introduction

	1. The rationale of conducting the present study is not well-justified. Authors cited some previous studies indicating the effectiveness and benefits of FSL and simply concluded that the existing literature is of limited scientific validity. Not sure how authors derived this conclusion. 2. The description of FSL should be moved to the Methods section. Instead, authors should provide a better overview (literature review) regarding the background of self-monitoring blood glucose. Methods 1. Given that the study was carried out in 13 public hospitals, n=165 participants recruited seems low. 2. Also, it is confusing that why authors needed a sample size calculation given that the present study design is framed as an observational study. Authors would need to specify the correct study design first. 3. If the study design is observational, how the authors defined the study baseline? 4. Authors would need to specify the data sources in terms of different outcome measures. For example, there are self-reported data, such as EQ-5D-Y, knowledge about diabetes treatment, satisfaction with treatment, safety, device adherence versus data obtained from EHR (HbA1c healthcare utilization, or pubertal stage). It is not clear how authors obtained the self-reported data. 5. For the cost estimate, usually, the intervention costs refer to the intervention program itself (in this case, the FSI). To gauge whether authors had reported the intervention costs properly, they would need to specify the process and procedure of receiving the FSL intervention first, including what resources and follow-up may be needed and what personnel may be involved.

REVIEWER	Palacios , Alfredo Institute for Clinical Effectiveness and Health Policy, Health Economics
REVIEW RETURNED	11-Apr-2023

GENERAL COMMENTS	*General comments: This study presents valuable "real-world" evidence regarding the use of the FreeStyle Libre (FSL) Glucose Monitoring System for children and adolescents with type 1 diabetes mellitus (T1DM) in Spain. However, four main issues should be addressed before the manuscript can be considered for publication: 1- Be explicit about the specific population and subpopulation when mentioning the results throughout the manuscript (abstract, introduction, discussion, and possibly the title). For example, the population consists of patients requiring more than six fingersticks per day, and the main result pertains only to patients with poor baseline control. This information should be explicitly presented throughout the manuscript. 2- The study aims to "inform health policy decision-making at the national level about coverage and public funding in these population groups." However, there is no discussion of the actual coverage decisions for these patients or the implications of this study for informing coverage decisions at the local level. 3- The authors need to address several important concerns about the costing methodology (which are detailed below).

	4- Given the study design, the authors should exercise caution when interpreting the results and moderate their conclusions throughout the manuscript (including the Introduction). *Specific comments: Abstract 1- Results: Please specify the subgroup size with baseline HbA1c >= 7.5%. Also, clarify whether the results related to reducing severe hypoglycemic events (SHE) correspond to self-reported SHE. 2- Conclusions: Reorder the sentences in the conclusion to prioritize the primary outcome (i.e., begin with the second sentence). "How this study might affect research, practice, or policy" section 3- Reorder the sentences to prioritize the primary outcome and moderate your interpretations. Based on your study design, the results cannot be considered "causal" evidence. Introduction 4- For international readers, add a paragraph describing the main characteristics of the Spanish health system/subsystems, including funding and covered population. 5- Include a paragraph discussing the trial evidence for the effectiveness of FSL for T1DM in children and young patients. 6- Most importantly, describe the coverage decisions/policies for FSL for T1DM children and adolescent patients in other countries and Spain. Materials and methods 7- Setting, logistics, and recruitment: Was the protocol published? If so, where can it be accessed? 8- Clarify the involvement of representatives and patient associations. Who were the representatives, and which patient associations were involved? 9- Specify which hospitals were included in the study and whether they can be considered "average" hospitals in Spain in terms of size, teaching status, etc. 10- Sample size calculation: Explain the rationale for requiring a sample size that is four times larger. 11- Statistical analysis: Address the issue of losing a significant number of patients during the follow-up, and provide a detailed analysis. For example, are there baseline differences between lost patients and those who continued? Discuss this issue in detail. 12- Why were two different software programs used for statistical analysis?

	13- Cost estimation (this is important): Please include all FSL-related costs (sensors, reader, etc.) for the base-case, "real-life" coverage analysis. Consider using different price/discount scenarios and account for acquisition costs, discontinuation costs, and adverse event costs. 14- Social costs: Clarify how absenteeism days were calculated, and provide all the inputs and relevant details for this analysis. Results 15- Table 2: Add the N for each sample/sub-sample column. 16- Move Tables 4 and 5 to the supplementary material, and bring at least Figure A1 into the main manuscript. Discussion 17- Regarding the statement "Two meta-analyses of case series on the effectiveness of the FSL yielded statistically significant HbA1c reductions in children/adolescents of -0.54% (n=447) [31] and -0.29% (n=959) [32], although with high statistical heterogeneity": Please clarify what you mean by "statistically significant with high statistical heterogeneity." 18- Additionally, specify whether these studies involved populations with poor metabolic control. If so, explicitly mention this. 19- Limitations: Discuss in detail the main implications of these limitations, such as the fact that the results cannot be interpreted as causal effects of FSL. 20- Add a paragraph about the implications of this study for coverage decisions. Reference studies conducted in other countries have discussed FSL coverage policy issues for the T1DM population. 21- Regarding the statement "However, the extremely low number of total hospitalizations during the monitoring study indicates that including this cost in the estimate would not have produced substantial changes in the results": Consider including all relevant costs for decision-makers and third-party payers, as it is essential to provide a comprehensive analysis. Conclusion 22- Moderate your conclusions and interpret the results conservatively to ensure that the study's findings are accurately represented.

REVIEWER	Svensson, Jannet Copenhagen University Hospital, Paediatric Department E
REVIEW RETURNED	30-Apr-2023

GENERAL COMMENTS	This is an observational study of 156 patients followed from starting using flash Glucose Libre system (FGL) and 12 months thereafter. There is no comparison group, but they should not be

	on FGL before initiation and should have had diabetes for at least 1 year. All participants were between 4 and 17 years. The study design has one major problem that is the primary outcome is likely to change as there is a known increase in HbA1c with age and diabetes duration and a known decrease when there are changes in treatment target. This means that a reduction is a sign of improvement – but if there has been focus on treatment target for HbA1c or time-in-range during the observation time reduction may reflects something else. The same is if more at starting pump treatment. If we assume that other factors with a possible positive effect are kept stable – then an improvement in glycemic parameters is not the most likely outcome and we may translate the reduction in HbA1c as a positive outcome. Even a stable HbA1c would in some cases could be interpret as a positive results! The study is design to look at numerous other outcomes both within mental health and health economics. Again, here the major drawback is the lack of a group for comparison since also some of these may change over a year when some of the children are entering puberty etc. So, I think there is a lot of good investigations and nice results, but there is not really that much new in showing an improvement in HbA1c and hypoglycemia! The perfect design that could provide evidence should either be randomized or quasi-eksperimentel where different centers are compared with different implementation strategies but then all treated at these clinics eligible for FGL should be screened for all the different factors included. The statistics seems appropriate for the design although I would recommend excluding some of the comparisons. It does not make sense to divide in two age groups and then test if age is different – the same goes for BMI (which should change between ages) and comparing mean HbA1c in two groups of HbA1c (< 7.5 and > 7.5) They should exclude these – they are NA and think of using an adjusted p value for multiple comparison. They also touch upon health economics, which is good, since it is important to show if new equipment is cost-beneficial. But I'm not sure I understand how they have estimated use of lancets and strips prior to starting FGL and the recommendation on when to control with a fingerprick is missing. Looking at the reduction it seems that there is still a relatively high use of fingerprick – more than I would expect – but it could be according to local guidelines. Again, the lack of a control group or a valid comparison group makes it difficult to use the estimates. In conclusion, the paper includes new parameters into the observational design and thereby contributes with new information. Regarding clear evidence the study should be designed to include some valid comparison which is not the individual patients – then they should have made a cross-over design. So they do not provide better evidence than previous studies but add to studies with this type of knowledge.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Tzeyu Michaud, University of Nebraska Medical Center, University of Nebraska Medical Center

Comments to the Author:

The manuscript sought to examine the effectiveness, safety and costs of FSL glucose monitoring system among children and adolescent with T1M. While this study may have some merits, the study design is not clear, lacking of sufficient details in a structured/clear format. For example, it is clear either in the Abstract or in the Methods section whether authors recruited participants who were already using FSL (thus the observation study) or recruited them to participate in the FSL intervention (thus a pre-post study design). In addition, there are significant amounts of grammar issues. Authors may benefit from the assistance of scientific writing service. Below are some comments (but not all), as it is not worth the effort to review the Results and Discussion sections without addressing the aforementioned issues first.

Response: Thank you for your observation. This has enabled us to realize that we have not been sufficiently clear with the study design nor the kind of patients to include. To improve we have made the following changes:
1. Study Design section in the abstract and methods: we have defined the study as prospective multicentre pre-post
2. We have modified the subjects section to make it clearer in the inclusion criteria that subjects did not have a prior FGM.

Abstract.
The abstract section should provide sufficient details for readers with an overview of the study without further looking into information in the main text. Apparently, authors have failed to do so. Some examples include, but not limited to:

1. The reason why the number of n=165 recruited participants was reduced to n=156 participants included in the analysis was not clear.

Response: In the results section of the main text, we explain that 9 patients were erroneously included since they did not fulfil inclusion criteria. In the abstract, we changed the text to “A total of 156 patients that met the inclusion criteria were included in the study”.

2. Need more details in terms of how costs were measured (instead of just the data source) and what are included.

Response: Thank you for your observation. We have added further details about the costs analysis in the abstract section of the reviewed manuscript.

3. Authors should align the outcomes reported in the Results section with the Outcome measures section. Several outcomes were not specified their operationalization, such as server hypoglycemic events, sensor usage time and scan, device adherence.

Response: We have now included all the outcomes in the two sections of the abstract. The operationalization of severe hypoglycaemic events or the questionnaires used are not included in the abstract due to the word limit. Sensor usage time and number of scans are directly interpretable.

Introduction

1. The rationale of conducting the present study is not well-justified. Authors cited some previous studies indicating the effectiveness and benefits of FSL and simply concluded that the existing literature is of limited scientific validity. Not sure how authors derived this conclusion.

Response: We have modified the introduction and improved the description of the current evidence on the effectiveness of FSL, including new references for systematic reviews and meta-analyses.

2. The description of FSL should be moved to the Methods section. Instead, authors should provide a better overview (literature review) regarding the background of self-monitoring blood glucose.

Response: According to your recommendations the paragraph reporting the technology has been moved to methods. As discussed above, new references have been included.

Methods

1. Given that the study was carried out in 13 public hospitals, n=165 participants recruited seems low.

Response: The inclusion criteria was that children had not used the technology previously. When the Spanish Ministry of Health identified the need to evaluate this technology only a few autonomous communities were offering it in their hospitals. Nonetheless, at the time of performing the study its use had been generalized, which reduced the number of hospitals taking part and subjects who fulfilled the study's inclusion criteria.

2. Also, it is confusing that why authors needed a sample size calculation given that the present study design is framed as an observational study. Authors would need to specify the correct study design first.

Response: Thank you for your observation, we have made the change in the study design.

3. If the study design is observational, how the authors defined the study baseline?

Response: Thank you for your observation. We have changed the study design to a pre-post study. This is a prospective multicentre pre-post study, whereby basal measurement of subjects were collected at the first consultation with the healthcare professional, before starting to use the FSL.

4. Authors would need to specify the data sources in terms of different outcome measures. For example, there are self-reported data, such as EQ-5D-Y, knowledge about diabetes treatment, satisfaction with treatment, safety, device adherence versus data obtained from EHR (HbA1c healthcare utilization, or pubertal stage). It is not clear how authors obtained the self-reported data.

Response: Thank you for your observation and suggestion. In our study, we use a combination of different data sources to evaluate multiple outcomes measurements. To measure quality of life, knowledge on diabetes, satisfaction with the treatment, safety and adherence to the device, self-informed data were used by subjects. These data were obtained by means of questionnaires performed on subjects or carers in the different follow up consultations with the health professional. Aside from self-informed data, we also used data obtained from electronic health records (EHR) to measure variables such as level of HbA1c, severe hypoglycaemia events, use of medical treatment

services or pubertal stage. The wording in the section “Setting, logistics and recruitment” and “Outcomes” has been improved.

5. For the cost estimate, usually, the intervention costs refer to the intervention program itself (in this case, the FSI). To gauge whether authors had reported the intervention costs properly, they would need to specify the process and procedure of receiving the FSL intervention first, including what resources and follow-up may be needed and what personnel may be involved.

Response: Thank you for your observation. The economic analysis is a secondary and additional endpoint to the study’s primary endpoint, whereby only some data available have been collated in the study on the use of resources (health and non-health). The collation of information did not purport to be exhaustive. Moreover, the timeline of the costs analysis is 12 months, which coincides with the duration of the main study, whereby the costs of follow up have not been considered. These limitations are highlighted in the revised discussion section of the manuscript.

Reviewer: 2
Dr. Alfredo Palacios , Institute for Clinical Effectiveness and Health Policy, University of York Centre for Health Economics
Comments to the Author:
*General comments:

This study presents valuable "real-world" evidence regarding the use of the FreeStyle Libre (FSL) Glucose Monitoring System for children and adolescents with type 1 diabetes mellitus (T1DM) in Spain. However, four main issues should be addressed before the manuscript can be considered for publication:

1- Be explicit about the specific population and subpopulation when mentioning the results throughout the manuscript (abstract, introduction, discussion, and possibly the title). For example, the population consists of patients requiring more than six fingersticks per day, and the main result pertains only to patients with poor baseline
control. This information should be explicitly presented throughout the manuscript.

Response: Thank you very much for your observation. The population has been explicitly included in the drawing up of the manuscript. The inclusion and exclusion criteria are specified in the “Participants” section.

2- The study aims to "inform health policy decision-making at the national level about coverage and public funding in these population groups." However, there is no discussion of the actual coverage decisions for these patients or the implications of this study for informing coverage decisions at the local level.

Response: Thank you for your observation. The information in this regard has been included. “Based on these results and other information sources (i.e., global research and clinical experts advice), the Spanish Ministry of Health decided to reimburse the FSL to T1DM aged 4-17 years-old, who undertake intensive therapy with insulin (multiple daily doses or with insulin pump), and require undertaking at least six finger pricks a day to self-monitor blood glucose.”

3- The authors need to address several important concerns about the costing methodology (which are detailed below).

Response: Thank you for your observation. The economic analysis is a secondary and additional endpoint to the study's primary endpoint, whereby only some data available have been collated in the study on the use of resources (health and non-health). The collation of information did not purport to be exhaustive. These limitations are highlighted in the revised discussion section of the manuscript.

4- Given the study design, the authors should exercise caution when interpreting the results and moderate their conclusions throughout the manuscript (including the Introduction).

Response: According to the suggestion of this and other reviewers, the study's limitations and conclusions have been reworded to highlight that given the study design a causal effect cannot be set out.

*Specific comments:

Abstract

1- Results: Please specify the subgroup size with baseline HbA1c >= 7.5%. Also, clarify whether the results related to reducing severe hypoglycemic events (SHE) correspond to self-reported SHE.

Response: We have now included the sample size of the two subgroups, and clarified that the reduction was observed on self-reported SHE.

2- Conclusions: Reorder the sentences in the conclusion to prioritize the primary outcome (i.e., begin with the second sentence). "How this study might affect research, practice, or policy" section.

Response: According to the editor's suggestions, the section "How this study might affect research, practice or policy" has been removed and replaced by a subsection "Strengths and limitations of this study".

3- Reorder the sentences to prioritize the primary outcome and moderate your interpretations. Based on your study design, the results cannot be considered "causal" evidence.

Response: Thank you for your observation, the study's conclusions have been reworded.

Introduction

4- For international readers, add a paragraph describing the main characteristics of the Spanish health system/subsystems, including funding and covered population.

Response: Thank you for your observation, this information has been included in the introduction.

5- Include a paragraph discussing the trial evidence for the effectiveness of FSL for T1DM in children and young patients.

Response: Thank you for your observation, further information in this regard has been included.

6- Most importantly, describe the coverage decisions/policies for FSL for T1DM children and adolescent patients in other countries and Spain.

Response: Thank you for your observation. We have sought information on the situation in other countries but we have not found scientific literature in this regard. Only journalistic news in regard to some countries. And according to what we found the device is partially or fully financed in over 30 countries.

Materials and methods

7- Setting, logistics, and recruitment: Was the protocol published? If so, where can it be accessed?

Response: If a protocol was set out in which they collaborated, aside from experts in health technologies evaluation, clinical experts, patient representatives and associations, and industry. However, as this was a Spanish Ministry internal document the protocol was not published.

8- Clarify the involvement of representatives and patient associations. Who were the representatives, and which patient associations were involved?

Response: In the protocol experts in the evaluation of health technologies,
clinical experts representatives of the Spanish Society of Internal Medicine (SEMI), Spanish Society of Diabetes (SED), Spanish Society of Endocrinology and Nutrition (SEEN), Spanish Society of Paediatric Endocrinology (SEEP), as well as patient representatives members of the Spanish Federation of Diabetes (FEDE), took part as authors and reviewers.

9- Specify which hospitals were included in the study and whether they can be considered "average" hospitals in Spain in terms of size, teaching status, etc.

Response: Thank you for your observation. In Appendix 1 we included sites taking part and cases included in the study. Moreover, the information on hospitals being reference hospitals on a national level; of the 13 hospitals taking part in the study, 6 are reference hospitals in Spain.

10- Sample size calculation: Explain the rationale for requiring a sample size that is four times larger.

Response: In the statistical section, we commented that we were interested in the effect of the intervention on subgroups defined by their baseline HbA1c level and age (four subgroups). If the difference in the effect size between the corresponding subgroups is small, the required sample size may be quite large; for instance, with balanced subgroups and an effect of the interaction half the value of the main effect, the sample required is 16 times larger than that needed to estimate the main effect to attain 80% power. We aimed to multiply the sample at least by 4 to increase the statistical power of the interactions as much as possible.

11- Statistical analysis: Address the issue of losing a significant number of patients during the follow-up, and provide a detailed analysis. For example, are there baseline differences between lost patients and those who continued? Discuss this issue in detail.

Response: Thank you for your observation. Due to the word limit the data imputation procedure was summarized. However, a comparability analysis of subjects lost during follow up and those that continued was made. This information has been included in the statistical analysis section and further details are given in Appendix 1.

12- Why were two different software programs used for statistical analysis?

Response: SPSS was used for the data cleaning part and baseline characteristics. STATA was used to perform the imputation and remainder of the analysis, which could not be performed in SPSS.

13- Cost estimation (this is important): Please include all FSL-related costs (sensors, reader, etc.) for the base-case, "real-life" coverage analysis. Consider using different price/discount scenarios and account for acquisition costs, discontinuation costs, and adverse event costs.

Response: Thank you very much for your comment.
As we discussed earlier, the economic analysis is a secondary endpoint and complementary to the primary endpoint. The start-up of a monitoring study has been used to collect data on the use of resources and make initial estimates of the cost of the intervention. Hence the limitations noted by the reviewer. We have tried to further detail cost analysis limitations in the discussion section of the revised manuscript.

14- Social costs: Clarify how absenteeism days were calculated, and provide all the inputs and relevant details for this analysis.

Response: Thank you for your comment. Absenteeism from work (number of days the caregiver was absent from work due to problems related to the child's T1DM) was collected at baseline and at 12 months. This information was multiplied by the cost per day of absenteeism estimated for Spain. To estimate this (cost per day of absenteeism in Spain), the cost per hour worked in Spain published by the Statistical Office of the European Union (Eurostat) has multiplied by the average number of daily working hours worked in Spain published in the Labour Force Survey (LFS) of the INE (Spanish Office of National Statistics).

Results
15- Table 2: Add the N for each sample/sub-sample column.

Response: Thank you very much for the suggestion, the table has been modified.

16- Move Tables 4 and 5 to the supplementary material, and bring at least Figure A1 into the main manuscript.

Response: Thank you for your observation. Given that the journal format enables us to include up to 5 tables and they are not large, we will leave them in the body of the manuscript. Figure 1 was included in the journal platform to be attached in the manuscript body. This has been included so that it can be seen without any problems.

Discussion
17- Regarding the statement "Two meta-analyses of case series on the effectiveness of the FSL yielded statistically significant HbA1c reductions in children/adolescents of -0.54% (n=447) [31] and -0.29% (n=959) [32], although with high statistical heterogeneity": Please clarify what you mean by "statistically significant with high statistical heterogeneity."

Response: It means that the pooled average effect was significant, but there was high statistical heterogeneity, that is, the observed effect varied widely across studies, and therefore, the certainty about the true effect diminishes. We have changed the wording “…although the effect was highly variable across studies”.

18- Additionally, specify whether these studies involved populations with poor metabolic control. If so, explicitly mention this.

Response: We have now included this explicitly, with the range of baseline HbA1c values.

19- Limitations: Discuss in detail the main implications of these limitations, such as the fact that the results cannot be interpreted as causal effects of FSL.

Response: Thank you for your observation, the suggested corrections have been made.

20- Add a paragraph about the implications of this study for coverage decisions. Reference studies conducted in other countries have discussed FSL coverage policy issues for the T1DM population.

Response: Thank you for your observation, the decision has been included on a national level in the manuscript's discussion. In reference to this and another prior observation, we sought information on the situation in other countries. However, we did not find scientific literature in this regard, only journalistic news in regard to some countries. And according to that found the device is financed partially or fully in over 30 countries. However, as discussed, there is no scientific report of these decisions.

21- Regarding the statement "However, the extremely low number of total hospitalizations during the monitoring study indicates that including this cost in the estimate would not have produced substantial changes in the results": Consider including all relevant costs for decision-makers and third-party payers, as it is essential to provide a comprehensive analysis.

Response: As previously mentioned, the economic analysis is a secondary endpoint and complementary to the primary endpoint. The start-up of a monitoring study has been used to collect data on the use of resources and make initial estimates of the cost of the intervention. Unfortunately, we do not have enough information to expand this initial analysis.

Conclusion

22- Moderate your conclusions and interpret the results conservatively to ensure that the study's findings are accurately represented.

Response: Thank you for your observation, the conclusions have been modified according to your recommendations and that of other reviewers.

Reviewer: 3

Dr. Jannet Svensson, Copenhagen University Hospital

Comments to the Author:

This is an observational study of 156 patients followed from starting using flash Glucose Libre system (FGL) and 12 months thereafter. There is no comparison group, but they should not be on FGL before initiation and should have had diabetes for at least 1 year.

All participants were between 4 and 17 years.

The study design has one major problem that is the primary outcome is likely to change as there is a known increase in HbA1c with age and diabetes duration and a known decrease when there are changes in treatment target. This means that a reduction is a sign of improvement – but if there has been focus on treatment target for HbA1c or time-in-range during the observation time reduction may reflects something else. The same is if more at starting pump treatment. If we assume that other factors with a possible positive effect are kept stable – then an improvement in glycemic parameters is not the most likely outcome and we may translate the reduction in HbA1c as a positive outcome. Even a stable HbA1c would in some cases could be interpret as a positive results!

The study is design to look at numerous other outcomes both wihin mental health and health economics. Again, here the major drawback is the lack of a group for comparison since also some of these may change over a year when some of the children are entering puberty etc.

So, I think there is a lot of good investigations and nice results, but there is not really that much new in showing an improvement in HbA1c and hypoglycemia! The perfect design that could provide evidence

should either be randomized or quasi-experimental where different centers are compared with different implementation strategies but then all treated at these clinics eligible for FGL should be screened for all the different factors included.

The statistics seems appropriate for the design although I would recommend excluding some of the comparisons. It does not make sense to divide in two age groups and then test if age is different – the same goes for BMI (which should change between ages) and comparing mean HbA1c in two groups of HbA1c (<7.5 and >7.5)
They should exclude these – they are NA and think of using an adjusted p value for multiple comparison.

Response: Thank you for your suggestions. The suggested changes have been made.

They also touch upon health economics, which is good, since it is important to show if new equipment is cost-beneficial. But I'm not sure I understand how they have estimated use of lancets and strips prior to starting FGL and the recommendation on when to control with a fingerprick is missing. Looking at the reduction it seems that there is still a relatively high use of fingerprick – more than I would expect – but it could be according to local guidelines. Again, the lack of a control group or a valid comparison group makes it difficult to use the estimates.

Response: Thank you very much for your comment. We agree with the reviewer that the main limitation of this study lies in its uncontrolled design, which precludes comparison with an untreated group. Regarding the economic analysis, we would like to clarify that it is a secondary endpoint and complementary to the primary endpoint. The start-up of a monitoring study has been used to collect data on the use of resources and make initial estimates of the cost of the intervention. The use of test strips and lancets was collected at baseline (without FSL) and at 12 months (using FSL). We believe that the significant use of these resources is due to the short learning time on the use of the device, on both the doctor and patient side.

In conclusion, the paper includes new parameters into the observational design and thereby contributes with new information. Regarding clear evidence the study should be designed to include some valid comparison which is not the individual patients – then they should have made a cross-over design. So they do not provide better evidence than previous studies but add to studies with this type of knowledge.

Response: Indeed, as the reviewer points out, the study's uncontrolled design precludes causal inferences and results from comparative studies are needed to draw definitive conclusions. We suppose that treatment targets for HbA1c were not changed by the introduction of the FSL, but we cannot be sure of this. Even if they were not changed, a "novelty effect" related to the use of a technological device could introduce a motivation bias that could affect self-management habits.
We have now discussed these issues in the limitations section.

VERSION 2 – REVIEW

REVIEWER	
REVIEW RETURNED	

GENERAL COMMENTS	

REVIEWER	
REVIEW RETURNED	

GENERAL COMMENTS	

REVIEWER	
REVIEW RETURNED	

GENERAL COMMENTS	

REVIEWER	
REVIEW RETURNED	

GENERAL COMMENTS	

REVIEWER	
REVIEW RETURNED	

GENERAL COMMENTS	

VERSION 2 – AUTHOR RESPONSE

VERSION 3 – REVIEW

REVIEWER	Palacios , Alfredo Institute for Clinical Effectiveness and Health Policy, Health Economics
REVIEW RETURNED	23-Aug-2023

GENERAL COMMENTS	I would like to thank the authors for taking the time to consider and incorporate my suggestions into the revised manuscript. The work that has been done in response to the reviewers' comments is commendable and has helped to enhance the paper's quality. I appreciate the efforts made in this revision, and I believe the manuscript is now suitable for publication. Thank you for addressing these concerns.

REVIEWER	Svensson, Jannet Copenhagen University Hospital, Paediatric Department E
REVIEW RETURNED	03-Sep-2023

GENERAL COMMENTS	Effectiveness, safety and costs of the FreeStyle Libre Glucose Monitoring System for children and adolescents with T1DM in Spain: a prospective uncontrolled pre-post study By Himar González-Pacheco et al. This is a pre-post intervention study – where the implications for HbA1c and the costs when FreeStyle Libre® is initiated are evaluated. The study design is still not optimal since they don't have a comparison besides each individuals own pre-values to post intervention. This has been better addressed and included in the discussion as limitation than in the previous version of the paper. General comments:

	I still think there is limited new information, and to many analyses given the power and the lack of using some of these information in a proper discussion. I would still recommend removing some of the analysis and focus the paper more on the important outcomes such as HbA1c and costs. Especially a lot of the interaction analyses are of low value and underpowered and perhaps not that important unless it is used for a discussion of how to improve scans or use of device in older adolescents and those with higher HbA1c. The test of interaction with HbA1c > 7.5 should also include the importance of "regression towards the mean" which means we would expect purely based on this phenomenon that the group with low HbA1c will tend to increase and the group with high hbA1c would tend to decrease. Regarding costs – I don't understand why the costs of the Free style libre is omitted – it should be included even if the sensors in this study was provided free of charge – in future health costs it is and extra costs to monitoring. You do have the costs from the clinic since many al ready are treated with Free Style libre! Specific comments: There are some repetitions from introduction to discussion – perhaps some of the introduction could be reduced and elaborated more in the discussion? The authors refer in the introduction to the meta-analysis and write "but in this case most studies had an uncontrolled design". Unfortunately, their study is also uncontrolled - so it seems a bit strange to point this out when they themselves do not change this. How do they add to current knowledge by repeating an uncontrolled design? The argument that the previous studies "is of limited scientific validity" is not really changed by this study – so the argument for following through with this study should be very focused on where this study is better or addresses new areas. Perhaps saying evidence is scarce or limited in children is better? "The interaction effect with baseline HbA1c level was not statistically significant for these two variables (P=0.117 and P=0.108, respectively). However, these analyses were underpowered and the descriptive statistics suggest different subgroup effects, although none was statistically significant". – so you can't really say anything about – then why do this analysis and include it in the paper? Figure A1: include costs of free style libre either as separate or part of monitoring In many of the tables there are 156 participants – but in the follow-up on skin issues there are only 128. It should be clear if the numbers in the tables only refer to baseline and not end of follow-up. "This interpretation is speculative since the commented results on self-reported SH were not statistically significant and underpowered, but it would help account for the unexpected significant worsening in self-perceived general health observed in the subgroup of poor baseline HbA1c monitoring." Or perhaps a

	more realistic interpretation is regression towards the mean and naturally increase in HbA1c with age? What is meant by “correct HbA1c” – what would be an incorrect measure of HbA1c? Should it be high and low HbA1c? The interaction with age – is not significant – nor is the descriptive statistics? As mentioned before there is a problem with the suspected increase over time – and more focus on the possible drivers such as less use of the equipment (usage time and scans) could be interesting. As well the lack of increase in knowledge – although these questions seems to be parents knowledge about insulin pump and nothing about the sensor – so perhaps that is not the relevant questionnaire to monitor knowledge when a Free Style libre is added? This should be addressed! Some minor errors: HbA1c monitoring is correct word to use – should it be HbA1c outcome? B=-1.27 (-1.89 to 0.65; P<0.001) – should be -0.65 or it is not significant?

VERSION 3 – AUTHOR RESPONSE

Reviewer:2
Dr. Alfredo Palacios , Institute for Clinical Effectiveness and Health Policy, University of York Centre for Health Economics

Comments to the Author:

1. I would like to thank the authors for taking the time to consider and incorporate my suggestions into the revised manuscript. The work that has been done in response to the reviewers' comments is commendable and has helped to enhance the paper's quality. I appreciate the efforts made in this revision, and I believe the manuscript is now suitable for publication. Thank you for addressing these concerns.
Response: Thank you very much for your comment. Thanks to your suggestions and those of the other reviewers, we have been able to improve our publication.

Reviewer: 3
Dr. Jannet Svensson, Copenhagen University Hospital
Comments to the Author:
Effectiveness, safety and costs of the FreeStyle Libre Glucose Monitoring System for children and adolescents with T1DM in Spain: a prospective uncontrolled pre-post study
By Himar González-Pacheco et al.
1. This is a pre-post intervention study – where the implications for HbA1c and the costs when FreeStyle Libre® is initiated are evaluated. The study design is still not optimal since they don’t have

a comparison besides each individuals own pre-values to post intervention.
This has been better addressed and included in the discussion as limitation than in the previous version of the paper.
Response: Thank you very much for your comment. Indeed, not having a comparator group is a significant limitation of this study, and we have wanted to reflect this in the publication.

General comments:
1. I still think there is limited new information, and to many analyses given the power and the lack of using some of these information in a proper discussion. I would still recommend removing some of the analysis and focus the paper more on the important outcomes such as HbA1c and costs. Especially a lot of the interaction analyses are of low value and underpowered and perhaps not that important unless it is used for a discussion of how to improve scans or use of device in older adolescents and those with higher HbA1c. The test of interaction with HbA1c > 7.5 should also include the importance of "regression towards the mean" which means we would expect purely based on this phenomenon that the group with low HbA1c will tend to increase and the group with high hbA1c would tend to decrease.

Response: We have included in the discussion that, given the uncontrolled nature of the study, the subgroup results based on their baseline HbA1c levels may simply reflect a regression to the mean effect. We have also noted that in the referenced meta-analysis on continuous glucose monitoring, which includes randomized controlled trials, a more significant effect was found in patients with poor baseline control, although this evidence also has its limitations. We have chosen to retain all analyses and variables. Given the study's limitations in terms of its design, as pointed out by the reviewer in their comments, we aim to provide information on a wide range of variables.

2. Regarding costs – I don't understand why the costs of the Free style libre is omitted – it should be included even if the sensors in this study was provided free of charge – in future health costs it is and extra costs to monitoring. You do have the costs from the clinic since many al ready are treated with Free Style libre!
Response: Thank you for the comment. We have included the unit cost of the sensor in the cost analysis.

Specific comments:

1. There are some repetitions from introduction to discussion – perhaps some of the introduction could be reduced and elaborated more in the discussion?
Response: The introduction is short and there is no extensive content repeated in the discussion. We just have shortened the statement in the introduction about the RCT on the effectivenes of FSL in adolescents, and we have now commented its result on satisfaction in the discussion.

2. The authors refer in the introduction to the meta-analysis and write "but in this case most studies had an uncontrolled design". Unfortunately, their study is also uncontrolled - so it seems a bit strange to point this out when they themselves do not change this. How do they add to current knowledge by repeating an uncontrolled design? The argument that the previous studies "is of limited scientific validity" is not really changed by this study – so the argument for following through with this study should be very focused on where this study is better or addresses new areas. Perhaps saying evidence is scarce or limited in children is better?

Response: In the mentioned paragraph, it is highlighted that, in contrast to T2DM, in patients with T1DM, most of the available evidence (including minors) is of an uncontrolled nature. This is necessary to inform the reader accurately about the quality of the available evidence, although it can certainly imply that this study aims to overcome that limitation, which is not the case. To try to avoid,

as much as possible, the text promoting such an interpretation, the order of the sentence providing this information has been changed, and the phrase referring to the 'limited scientific validity' of the evidence has been removed.
3. After presenting the available evidence in this paragraph, the following paragraph explains that the study was promoted by the Spanish Government to obtain evidence in the Spanish context (this clarification has been added to the text). Although the research team proposed a controlled design to obtain higher-quality evidence, legal and logistical requirements prevented it.

“The interaction effect with baseline HbA1c level was not statistically significant for these two variables (P=0.117 and P=0.108, respectively). However, these analyses were underpowered and the descriptive statistics suggest different subgroup effects, although none was statistically significant”. – so you can’t really say anything about – then why do this analysis and include it in the paper?

Response: The low statistical power and the lack of significance prevent us from drawing firm conclusions, but it does not preclude a hypothetical interpretation of the results, which should be confirmed in future higher-quality studies. For example, in the case of age, no results were statistically significant, but the point estimates of the differences in HbA1c and severe hypoglycemia show a better outcome in children than in adolescents, which is consistent with the better adherence to the device observed in the former. This improved adherence by younger participants has also been observed in other studies, both on flash monitoring and continuous glucose monitoring, as discussed in the paper.
In the case of severe hypoglycemia, even though, as suggested by the reviewer, the effect may simply reflect regression to the mean, we have chosen to retain all analyses. In patients with high baseline HbA1c levels, who significantly reduced their HbA1c levels, the percentage of participants with at least one event of severe hypoglycemia increased from 26% to 38%; although this difference did not reach statistical significance, we believe it deserves mention.

4. Figure A1: include costs of free style libre either as separate or part of monitoring
Response: Thank you for your comment. We have added the cost of the sensor to the figure.

5. In many of the tables there are 156 participants – but in the follow-up on skin issues there are only 128. It should be clear if the numbers in the tables only refer to baseline and not end of follow-up.
Response: The difference between the tables is due to the loss of information over the course of the study follow-up. The variable "Mild adverse effects caused by the sensor" was excluded from the imputation model due to its specific nature and the type of analysis intended for it, which aimed to assess the reduction of mild adverse effects caused by the sensor throughout the study.

6. “This interpretation is speculative since the commented results on self-reported SH were not statistically significant and underpowered, but it would help account for the unexpected significant worsening in self-perceived general health observed in the subgroup of poor baseline HbA1c monitoring.” Or perhaps a more realistic interpretation is regression towards the mean and naturally increase in HbA1c with age?
Response: Regression toward the mean cannot explain the quality of life outcome since the mean values for this variable at baseline were very similar across the subgroups of baseline HbA1c. We argue that if there has indeed been a deterioration in self-perceived general health in the subgroup with poorly controlled baseline HbA1c, this could be more relevant to overall health self-perception than the reduction in HbA1c, which may go unnoticed in a person's self-perception. We believe this is not an unrealistic interpretation, always considering the limitations of the study that have been previously discussed.

7. What is meant by “correct HbA1c” – what would be an incorrect measure of HbA1c? Should it be high and low HbA1c?

Response: Yes, we have changed that expresión to “controlled HbA1c levels”.

8. The interaction with age – is not significant – nor is the descriptive statistics? As mentioned before there is a problem with the suspected increase over time – and more focus on the possible drivers such as less use of the equipment (usage time and scans) could be interesting.
Response: As previously mentioned, while the results were not statistically significant, opposite trends were observed among age subgroups, which were related to lower (not poor) adherence to the device observed in adolescents, supported by existing literature on glucose monitoring in general. We have included a sentence discussing potential specific barriers for adolescents and young adults in using glucose sensors and the need to implement strategies tailored to this population, with new references.

9. As well the lack of increase in knowledge – although these questions seems to be parents knowledge about insulin pump and nothing about the sensor – so perhaps that is not the relevant questionnaire to monitor knowledge when a Free Style libre is added? This should be addressed!
Response: Our aim was not to assess knowledge of the FreeStyle Libre (FSL) system; we assumed that knowledge of it was initially low or nonexistent and that it improved after the intervention, as the device was used regularly and generally without significant operational problems or serious adverse effects (which could occur with improper use due to lack of knowledge). Our objective was to evaluate knowledge related to diabetes treatment (with insulin injections or an insulin pump) to ascertain whether the use of the FSL and the knowledge requirements associated with its use contribute to an overall improvement in knowledge about correct insulin treatment application. Therefore, we do not consider this to be a study limitation. Nevertheless, we have specified "knowledge of diabetes treatment" whenever the questionnaire is mentioned.

Some minor errors:

1. HbA1c monitoring is correct word to use – should it be HbA1c outcome?
B=-1.27 (-1.89 to 0.65; P<0.001) – should be -0.65 or it is not significant?

Response: Thank you for your comment; we have noticed this minor error. Indeed, it should be -0.65. It has been corrected in the body of the article and the table. Thank you very much.

VERSION 4 – REVIEW

REVIEWER	Svensson, Jannet Copenhagen University Hospital, Paediatric Department E
REVIEW RETURNED	26-Oct-2023

GENERAL COMMENTS	The paper have been improved although there is still a problem with the design to fully answer the question if Freestyle Libre is effective and describe costs. The discussion is better balanced. The argument for repeating a non randomized study is still vague. I acknowledge the decision to keep a lot of unpowered analysis in the paper - but don't agree it improves the paper since you can hardly say they are explorative. In the abstract they still conclude: "Using FSL could reduce healthcare and productivity losses related costs" - but cost in figure A1 show the costs are similar (or even a bit higher with

	FSL?)? I don't understand why this figure is in a supplement when it is the main outcome - and even presented in the article title? Small typo in table 3? 12 months (n=128) 14 (1.9) 6 (4.7) 2 (1.6) There are some Spanish words in supplemental material. No further comments.

VERSION 4 – AUTHOR RESPONSE

Reviewer: 3
Dr. Jannet Svensson, Copenhagen University Hospital

Comments to the Author:

1. The paper have been improved although there is still a problem with the design to fully answer the question if Freestyle Libre is effective and describe costs.

Response: We believe that in the final version of the article, it is explicitly acknowledged that the uncontrolled design is a fundamental limitation in drawing firm conclusions. However, the only alternative to this limitation is to forgo publishing this study, an option that we obviously do not endorse (and hope the editor doesn't either). In any case, we sincerely appreciate the reviewer's critical comments, which have undoubtedly helped us improve the article, within the aforementioned limitations.

2. The discussion is better balanced. However, the argument for repeating a non randomized study is still vague.

Response: We appreciate your acknowledgment of the improved balance in the discussion. As we mentioned in the study's limitations section, the primary limitation of this research is its uncontrolled design, which precludes comparison with an untreated group. We have recognized this limitation, and it has been discussed in the manuscript. We believe that the findings from this study still contribute valuable insights, despite this limitation.

3. I acknowledge the decision to keep a lot of unpowered analysis in the paper - but don't agree it improves the paper since you can hardly say they are explorative.

Response: We do not claim that the inclusion of these results represents an improvement of the article, but we still maintain the decision to provide interested readers with all available results.

4. In the abstract they still conclude: "Using FSL could reduce healthcare and productivity losses related costs" - but cost in figure A1 show the costs are similar (or even a bit higher with FSL?)? I don't understand why this figure is in a supplement when it is the main outcome - and even presented in the article title?

Response: We greatly appreciate your comment and we have added the figure in the main document according to your suggestion. Additionally, we have tried to be clearer in the abstract of the revised manuscript. Thank you very much.

5. Small typo in table 3?
12 months (n=128)
14 (1.9)
6 (4.7)
2 (1.6)

Response: Thank you very much for your comment. Table 3 has been reviewed and corrected.

6. There are some Spanish words in supplemental material.

Response: Thank you very much for your suggestions. The changes have been made.