PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Cost-effectiveness of home versus hospital management of children at onset of Type 1 Diabetes: The DECIDE randomised controlled trial
AUTHORS	McCarroll, Zoe; Townson, J; Pickles, Timothy; Gregory, John; Playle, Rebecca; Robling, Michael; Hughes, Dyfrig

VERSION 1 – REVIEW

REVIEWER	Wherrett, Diane SickKids Research Institute
REVIEW RETURNED	01-Oct-2020

GENERAL COMMENTS	This manuscript describes the economic evaluation of a clinical trial of home versus hospital care for the management of newly diagnosed type 1 diabetes in children (DECIDE RCT). The previously published results of this trial showed no significant difference in hemoglobin A1c (primary outcome) at 24 months. The authors included detailed analyses of direct and indirect healthcare costs and found that home management was less costly based on the increased cost of a three-day hospitalization at the time of initiation of insulin therapy in the hospitalization group. Costs over the remainder of the 24 month follow-up period were not different. The manuscript lays out the costs clearly and accounts for a large number of factors that appears to be very complete. It explains the reasons for choice of analyses and the limitations encountered well. The manuscript could be made more readable for a general audience with interest in pediatric diabetes but without expertise in economic analyses. Several terms should be explained such as cost effectiveness plane and cost-effectiveness acceptability curve. More detail could be added to the legend for Figure 1.

REVIEWER	Ramchandani , Neesha New York University
REVIEW RETURNED	04-Nov-2020

GENERAL COMMENTS	This is a very comprehensive and well-written manuscript addressing the cost of care for a child newly diagnosed with type 1 diabetes in the UK. While I thought I had read papers like this before, the authors address how their manuscript adds to the literature - this manuscript specifically looks at costs for the UK, where other papers did not, and it looks at indirect as well as direct costs. I had just a few comments for revisions.

	1. In your tables, I would like to see indications of the statistically significant differences, even if you just put an asterix (*) next to each one and list the p-value that it corresponds to at the bottom or in the title. 2. I would address that you did not evaluate time in range in addition to HbA1c because it was not something that clinicians were specifically assessing and documenting at the time the DECIDE study data were being collected. 3. I did not understand why you put at 3.5% discount rate after the first 12 months, even though you write it was recommended by NICE. Please consider adding one more sentence to this section to explain why, especially for the reader who is not familiar with the NHS or UK systems. Thank you for doing such a comprehensive study and writing it up so clearly!

REVIEWER	Marques, Elsa University of Bristol, Bristol Medical School
REVIEW RETURNED	03-Mar-2021

GENERAL COMMENTS	Overview: This is a very large and complex economic evaluation alongside a trial to improve initial training and management of T1DM in children. There have been several trials comparing similar treatments and showing that home management is cost-effective. In this respect, this trial or economic evaluation results are not new. I am not surprised that there is no evidence for a clinical effect; previous studies have not found one either. It is a shame the trial is not powered for equivalency. Ultimately what you would have liked is show that home training/management would be no worse than hospital management, and let the costs decide what would be the best use of societal resources. Since the trial is relatively small, we cannot jump to that conclusion. The evaluation is exhaustive and well conducted. The team has attempted to collect and cost an awful lot of detailed resources over 2 years. The evaluation methods are all tried and tested but basic. I do feel there is room for improvement on how this evaluation is described to the audience and suggest some minor points in my notes below. I find that reporting ICERs in this occasion are not very informative. For a trial that wanted to detect a minimum different of 5mmol/mol (SD 14) and observed a difference of 0.2 mmol/mol, this value is very close to zero when standardised. The uncertainty around the health benefit is so large (-6.282 to 5.695), that when the ICER is bootstrapped, the CIs are very wide (-73368.77 to 88236.77) rendering the mean ICER value uninformative, reflected in a 50/50 probability of the intervention being cost-effective. We cannot really make judgements on the cost-effectiveness of the intervention. The authors also acknowledged that we have no value judgement for cost per change in mmol/mol test for decision-making. The ICER results then, if derived, should be downplayed (e.g. not appear in the abstract or major findings sections) The trial results do not necessarily mean that there is no difference between arms, only that the trial was not able to detect one. Do not revert to a cost-minimisation study, but the evaluation could potentially become more comprehensive if it incorporated other health benefits (namely quality of life measures in children

	measured in the DECIDE trial) in a cost-consequences table. It does not look like the specific T1DM scale of the PedsQL used in DECIDE includes the generic QoL domains that could be mapped to utilities using Khan's 2014 algorithm. They could still be reported and/or discussed within this evaluation. Minor notes: Abstract: P2 line 19. Inconsistent with article where the authors state NHS+PSS as base case. No PSS resources were identified in this evaluation in the methods on p5, but the supplementary material tables include unit costs for social workers. P2 line 27. Consider abstaining from statements of significance. Study not powered to detect difference in costs. Introduction: The introduction could be shortened. For example, many details of the DECIDE trial are repeated in the methods, and could be erased from here. The authors could improve their referencing in the article. I have not checked all references, but on a quick glance of the ones in p3, I noted: - line 11/12. Not sure how reference 2 is evidence of growing emphasis. - lines 16/17. The authors reference two trials and a review (and describe more studies in the 3rd paragraph) that show how T1D can be cost-effectively managed outside the hospital, therefore disproving the point of the sentence. - line 36/37. The first sentence is just repeating a subjective statement made by the Swedish authors in their paper. It is a spurious statement. - lines 45/47. Swedish trial results were misstated; they were not powered on equivalence. "Absence of evidence is not evidence of absence" (Doug Altman). Methods: P4 lines50/52. Please explain what home supervisions/management consisted of, performed by whom, how, and for how long, without having to search for other papers. P5 Costs. There is some lack of clarity in the description of the cost categories included for analysis, e.g. not clear what medical equipment consisted of. Consider uploading data collection tools in the Dirum.org database. I don't understand the point of Table1 in supplementary materials. P5 lines 4/8. Unclear. If resource use prior to diagnosis was not included, which resource use was collected in CRFs at baseline? State somewhere when baseline took place. P5 lines 18/27. DHSC reference costs provide higher external validity of the findings; they are not necessarily second best. Please justify why authors favour local over national unit costs. P5 lines 39/40. Please state which approach you used for this (human capital/ friction costs, would they make a difference?). Did

	you consider (if appropriate) any other measures of informal care? E.g. care from other relatives? P5 lines 40/42. There are numerous ways to value children's missed days off school (e.g. Andronis et al 2019, https://pubmed.ncbi.nlm.nih.gov/31408769/). This approach takes the perspective of the school/government spending resources keeping a child in school when desk not being used. There are others. Please justify your choice. P5 lines 50/55. I agree a model would not be necessary because one arm is dominant in the short-term – home management – and you are assuming being managed at home initially would not lead to higher hospitalizations in the longer-term. I find the authors justification a bit counter-intuitive, and could be cut. In P6, Analytical methods, I was expecting this section to make statements on: a) was the analysis intention to treat? b) How missing cost and outcome data was dealt with and that the base case analysis a complete case analysis (I figure). c) Which regression analysis was performed? Were costs and outcomes jointly estimated? Is there a relationship between the two (possibly not if no difference in outcomes, but was this investigated?) P6 line 51/53. I am unclear why the authors needed to bootstrap costs and effects in Excel 1,000, when they had bootstrapped them in Stata 10,000 times. They can create the cost-effectiveness planes with the Stata replications, either directly in a Stata graph, or copy-pasting the bootstrapped values to excel and creating a graph from there. P6 lines 59/60. In a post-COVID scenario, would the intervention still be delivered in the same way? Would some of the home management be done online? And if so, this could perhaps be reflected in a scenario analysis in SA. Home management being the cheapest arm, it should not make a difference to the results, but could be discussed in the discussion. Results: Lines 8/14. Was there no incomplete cost data, at 24 months, for all cost categories? Suggest removing decimal points from cost tables, makes them easier to read. Table 3. I think the calculations in this table should be rechecked. I find the costs and effects with confidence intervals around the point estimates in Tables 14,15, 16 in supplementary materials more informative, and that information to could included in the main article tables. Discussion: I feel the discussion overplays the results of the economic evaluation somewhat and could be re-worked if the authors decide to take on some of the above suggestions on board.

	Thank you for the opportunity to review this very large and interesting evaluation, and all the best with the proceedings of this submission.

VERSION 1 – AUTHOR RESPONSE

Responses to reviewer's comments

Responses to reviewer's comments

Reviewer: 1	
Comments	Response
The manuscript could be made more readable for a general audience with interest in pediatric diabetes but without expertise in economic analyses. Several terms should be explained such as cost effectiveness plane and cost-effectiveness acceptability curve. More detail could be added to the legend for Figure 1.	We agree this is useful information to have, we have added the following two sentences to Methods page 6, last paragraph. "The cost-effectiveness plane is used to visually represent the differences in costs and health outcomes between arms in two dimensions." "The (cost-effectiveness acceptability curve) CEAC is used to summarise the impact of uncertainty on the result of an economic evaluation. It represents the probability of an intervention being cost-effective for any given value of the cost-effectiveness threshold."
Reviewer: 2	
1. In your tables, I would like to see indications of the statistically significant differences, even if you just put an asterix (*) next to each one and list the p-value that it corresponds to at the bottom or in the title.	These have been added to Table 1 and the relevant values have been added to Table 3 from the supplementary material. We have added in this manner, in order to follow CONSORT guidelines.
2. I would address that you did not evaluate time in range in addition to HbA1c because it was not something that clinicians were specifically assessing and documenting at the time the DECIDE study data were being collected.	We have added a sentence to the Discussion (Page 16) to address this point. "It should also be noted that at the time

	this study was conducted, few patients were using continuous glucose monitoring to allow us to collect data on 'time in range'.
3. I did not understand why you put at 3.5% discount rate after the first 12 months, even though you write it was recommended by NICE. Please consider adding one more sentence to this section to explain why, especially for the reader who is not familiar with the NHS or UK systems.	We have added a sentence to provide further explanation (last sentence Page 4). "We used this rate because all economic evaluations require that future costs and effects are discounted to present value to account for time preference. In the UK, the discount rate is set at 3.5% per annum."
Thank you for doing such a comprehensive study and writing it up so clearly!	Thank you for your comment it is much appreciated.
Reviewer: 3	
General comments	
I find that reporting ICERs in this occasion are not very informative. For a trial that wanted to detect a minimum different of 5mmol/mol (SD 14) and observed a difference of 0.2 mmol/mol, this value is very close to zero when standardised. The uncertainty around the health benefit is so large (-6.282 to 5.695), that when the ICER is bootstrapped, the CIs are very wide (-73368.77 to 88236.77) rendering the mean ICER value uninformative, reflected in a 50/50 probability of the intervention being cost-effective. We cannot really make judgements on the cost-effectiveness of the intervention. The authors also acknowledged that we have no value judgement for cost per change in mmol/mol test for decision-making. The ICER results then, if derived, should be downplayed (e.g. not appear in the abstract or major findings sections)	We acknowledge the limitation of HbA1c as the denominator of the ICER calculation. However, the ICER was the pre-specified primary outcome of the economic analysis, and it would therefore be inappropriate to not report it in the results section, based on its value. As it transpired, the ICER value has reduced importance given that home management dominated hospital management.
Do not revert to a cost-minimisation study, but the evaluation could potentially become more comprehensive if it incorporated other health benefits (namely quality of life measures in children measured in the DECIDE trial) in a cost-consequences table. It does not look like the specific T1DM scale of the PedsQL used in DECIDE includes the generic QoL domains that could be mapped to utilities using Khan's 2014 algorithm. They could still be reported and/or discussed within this evaluation.	We agree that it would be useful to assess the PedsQL as a potential way to estimate EQ-5D utilities. However, we didn't use the PedsQL™ generic core scales required, as detailed in the Khan 2014 paper, so unfortunately we are unable to add further analyses. Moreover, a search of the literature and the HERC database of mapping studies did not identify a suitable alternative.
Abstract	

1. P2 line 19. Inconsistent with article where the authors state NHS+PSS as base case. No PSS resources were identified in this evaluation in the methods on p5, but the supplementary material tables include unit costs for social workers.	Thank you for this comment, we've corrected this to NHS perspective.
2. P2 line 27. Consider abstaining from statements of significance. Study not powered to detect difference in costs.	We have revised the statement accordingly, by removing the word "significant".
Introduction	
3. The introduction could be shortened. For example, many details of the DECIDE trial are repeated in the methods, and could be erased from here.	We have removed detail in the Introduction where it is repeated in Methods. The introduction is now shorter.
4. The authors could improve their referencing in the article. I have not checked all references, but on a quick glance of the ones in p3, I noted: - line 11/12. Not sure how reference 2 is evidence of growing emphasis.	We agree, there was an error in the referencing, this has been corrected.
5. Lines 16/17. The authors reference two trials and a review (and describe more studies in the 3rd paragraph) that show how T1D can be cost-effectively managed outside the hospital, therefore disproving the point of the sentence.	We have revised these sentences to provide clarity, page 3.
6. Line 36/37. The first sentence is just repeating a subjective statement made by the Swedish authors in their paper. It is a spurious statement.	We agree, this sentence has been revised, page 3.
7. Lines 45/47. Swedish trial results were misstated; they were not powered on equivalence. "Absence of evidence is not evidence of absence" (Doug Altman).	We agree, this sentence has been revised, page 3.
Methods	
8. P4 lines50/52. Please explain what home supervisions/management consisted of, performed by whom, how, and for how long, without having to search for other papers.	We have provided additional information bottom of page 4
9. P5 Costs. There is some lack of clarity in the description of the cost categories included for analysis, e.g. not clear what medical equipment consisted of. Consider uploading data collection tools in the Dirum.org database. I don't understand the point of Table1 in supplementary materials.	All medical equipment related to managing Type 1 diabetes. We have added a sentence to page 5 for clarification, "Medical equipment included items such as testing strips, needles, and lancets." We have requested that the resource use questionnaire is upload to DIRUM. Table 1 in supplementary materials, describes the different data collection forms (CRFs), and the data collected at different time points. The CRFs are mentioned in some table footnotes, therefore we

	believe this table provides necessary clarity.
10. P5 lines 4/8. Unclear. If resource use prior to diagnosis was not included, which resource use was collected in CRFs at baseline? State somewhere when baseline took place.	We have added this sentence to Page 5 to provide clarity on baseline “Baseline data comprised of data collected from the day of diagnosis until day 3 of either home or hospital management.”
11. P5 lines 18/27. DHSC reference costs provide higher external validity of the findings; they are not necessarily second best. Please justify why authors favour local over national unit costs.	Most costs were nationally representative, coming from the NHS Reference costs, the British National Formulary, PSSRU compendium of unit costs, and the Drug Tariff. However, national costs lack granularity with regards to specific tests, hence our reliance on local figures. We have revised the sentence regarding local costs of hospital stays.
12. P5 lines 39/40. Please state which approach you used for this (human capital/ friction costs, would they make a difference?). Did you consider (if appropriate) any other measures of informal care? E.g. care from other relatives?	Time off work was costed based on national data on median weekly earnings (human capital approach). We agree that there are alternative approaches, which we might have considered in a sensitivity analysis. But as this was a secondary analysis, we felt it unnecessary to go to this level of detail. We did not consider the potential costs of care from other relatives.
13. P5 lines 40/42. There are numerous ways to value children’s missed days off school (e.g. Andronis et al 2019, https://eur03.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpubmed.ncbi.nlm.nih.gov%2F31408769%2F&data=04%7C01%7CTownson%40cardiff.ac.uk%7C45491d31ea8e4a17d20e08d8e7bd466d%7Cbdb74b3095684856bdbf06759778fcbc%7C1%7C0%7C637514147398904752%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=nmByhi2Co4IEd5fCJQ75Ocr5JvmcLUmNeT%2FjYzKTI6o%3D&reserved=0). This approach takes the perspective of the school/government spending resources keeping a child in school when desk not being used. There are others. Please justify your choice.	The difference in the number of days of schooling missed was very small between intervention groups (Supplementary Appendix table 12). Andronis’ paper provides an excellent overview of the topic, and the specific challenges. They found a “sizeable literature on time valuation methods in education, labour and transportation economics, much of this is not directly applicable to economic

	evaluation of health care interventions for children". We have revised the text in the Discussion accordingly: "There are a number of methodological challenges in assigning costs to days of missed schooling, with no clear consensus on the most appropriate approach [Ref Andronis]. We costed the time taken off school based on calculating an average cost spent per pupil per day, based on the Annual Report on Education Spending in England. [Ref 24] This may underestimate the economic consequences of forgone leisure time and educational achievement."
14. P5 lines 50/55. I agree a model would not be necessary because one arm is dominant in the short-term – home management – and you are assuming being managed at home initially would not lead to higher hospitalizations in the longer-term. I find the authors justification a bit counter-intuitive, and could be cut.	We believe that our justification is consistent.
15. In P6, Analytical methods, I was expecting this section to make statements on: a) was the analysis intention to treat? b) How missing cost and outcome data was dealt with and that the base case analysis a complete case analysis (I figure). c) Which regression analysis was performed? Were costs and outcomes jointly estimated? Is there a relationship between the two (possibly not if no difference in outcomes, but was this investigated?)	

VERSION 2 – REVIEW

REVIEWER	Ramchandani , Neesha New York University
REVIEW RETURNED	22-Apr-2021

GENERAL COMMENTS	This manuscript is very clear and well put together, and addresses a notable gap in the literature. The authors are also very forthright with their limitations and explaining why they chose to do what they did.

	There is 1 minor edit I would suggest, which can be addressed by the Editor. On pg. 16 of the entire document (pg. 15 of the revised manuscript), line 29, the word should be "affect", not "effect." Other than that, everything looks good!