

PEER REVIEW HISTORY

BMJ Paediatrics Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Growth Hormone treatment in Prader-Willi syndrome patients: systematic review and meta-analysis
AUTHORS	Passone, Caroline; Franco, Ruth; Ito, SImone; Trindade, Evelinda; Polak, Michel; Damiani, Durval; Bernardo, Wanderley

VERSION 1 – REVIEW

REVIEWER	Reviewer name: Sarah Nevitt Institution and Country: University of Liverpool United Kingdom Competing interests: I have no competing interests
REVIEW RETURNED	31-Jan-2020

GENERAL COMMENTS	<p>I have performed a statistical review of the manuscript "Growth Hormone treatment in Prader-Willi syndrome patients: a global systematic review and meta-analysis"</p> <p>The authors perform a systematic review and meta-analysis of randomised and non-randomised studies examining growth hormone treatment in Prader-Willi syndrome and conclude that growth hormone treatment results in an improvement of stature, body composition and BMI.</p> <p>I am unclear regarding certain elements of the methods employed:</p> <p>1) I do not understand exactly which study designs are eligible for this review. It is stated on page 4 that randomised and non-randomised (cohort and before and after studies) are eligible designs.</p> <p>However, it is also stated in the next paragraph that rhGH is to be compared to placebo or no drugs – this comparison cannot be made in before and after studies.</p> <p>Also, a quality assessment tool for case series is mentioned on page 5 – were case series also included?</p> <p>Further, within the results section it is stated that "For the meta-analysis process and for a better approach to all data, we transformed all cohort studies into before and after studies." I do not understand why or how this has been done.</p> <p>Please clarify the above points regarding the study designs</p> <p>2) I do not understand most of the content under the heading "Measures of treatment effect and units of analyses, " particularly the following sentences:</p> <p>"Different follow-up lengths were standardized by use of incidence densities of outcomes occurrence across patient-years followed.</p>
-------------------------	---

	<p>When different measures or instruments were used in the studies, standard errors were used to estimate outcomes”</p> <p>And also the formula to convert 25th and 75th percentiles. It seems unlikely a study would report a mean and percentiles (generally a median would be reported with percentiles).</p> <p>Please clarify and provide a reference for the formula.</p> <p>3) Please provide rationale for the subgroup analysis and how the cut off of 3.5 years was determined</p> <p>4) The authors are correct to perform separate meta-analyses of randomised and non-randomised studies but please clearly state this within the methods</p> <p>5) Please check the GRADE certainty judgements in Supplementary Files 5 and 6. For example, the analysis of Z-stature with serious risk of bias and very serious inconsistency should probably be graded lower than moderate.</p> <p>How were levels of association determined as ‘strong’ and ‘very strong’? Please provide some justification (ideally with a reference) for this.</p> <p>I also have some minor comments</p> <p>1) I don’t understand the use of the word ‘global’ in the title and what this refers to. I also suggest that the word ‘recombinant’ should be used in the title</p> <p>2) The terms ‘somatropin’ and ‘rhGH’ are used interchangeably. From my understanding, both terms refer to the same treatment but for consistency, I suggest choosing just one of these terms and using it throughout the manuscript</p> <p>3) Introduction: I suggest that the first paragraph of the introduction may need some more references. Does everything in this paragraph (prevalence rate, obesity, quality of life) come from reference 1 at the end of the paragraph?</p> <p>4) Methods: “All studies that had patients who received a subcutaneous application of rhGH at any dosage, ranging from 0.3 mg/m2/day to 1 mg/m2/day, were included.” I don’t understand this sentence – it says that ‘any dosage’ would be included but then quotes a range of doses? Or is ‘0.3 mg/m2/day to 1 mg/m2/day’ the range of dosages within the included studies? If so, I suggest stating this in the results section rather than the methods section.</p> <p>5) Methods, page 4: Are any of the listed outcomes primary outcomes of this review?</p> <p>6) Methods, page 4: Quality of life and behaviour ‘evaluated by different instruments.’ Presumably different validated instruments?</p> <p>7) Methods, page 5: “All articles PWS were evaluated independently by 2 reviewers.” PWS is not needed in this sentence</p> <p>8) Methods, page 5 (Selection of studies and data extraction): Please provide details of which outcomes are analysed as dichotomous variables and which are analysed as continuous variables</p> <p>9) Table 1 and Table 2: I’m not sure why ‘systematic review 2019’ is in the header of these tables?</p> <p>10) Table 1: Please provide details of the control groups in these studies. It would also be helpful to include the number of people in the rhGH and control groups in these tables.</p>
--	---

	<p>11) Results: “We performed the meta-analysis for the evaluation of growth hormone treatment versus placebo/no treatment in nine of the seventeen included RCTs” In the previous sentence, it is stated that 16 studies were included, not 17? Also, why could only 9 studies be included? Were appropriate data not available in the other studies?</p> <p>12) Results: “For the evaluation of cognitive function in patients with PWS, different questionnaires were used; therefore, global analysis was not possible.” I’m not sure what a ‘global’ analysis is, should this read ‘meta-analysis?’</p> <p>13) Assessment of heterogeneity and reporting bias, page 6: I-squared is a statistic rather than a test</p> <p>14) Assessment of heterogeneity and reporting bias, page 6: Is the p-value in this section relating to a test of the heterogeneity or of the treatment effect?</p> <p>15) Assessment of heterogeneity and reporting bias, page 6: “The possibility of publication bias was assessed by evaluating a funnel plot (Begger Test) for asymmetry” I cannot find a funnel plot within the Figures? Was there any evidence of publication bias?</p> <p>16) Results, page 9: When quoting results in the text please clearly state that the numbers in brackets are a confidence interval, for example 1.67 (95% CI 1.54-1.81) SDS. Please also clearly state in the text where heterogeneity is present.</p> <p>17) Results, page 9: “In the NRCT studies, the treated group had a significant improvement of stature (1.52 (0.86;2.16) SDS) during the 2,909 patients-years of follow-up (Figure 2).” Please check this result, the numbers on Figure 2 are different.</p> <p>18) Results, page 9: “We chose to report the growth velocity measure in centimetres and due to that one article that used SDS was excluded.” I do not understand what this means?</p> <p>19) Results, page 9: Lean mass – two studies seem to be included in each analysis but the text states three.</p> <p>20) Results, page 9-10: “The studies by Lindgreen 1998, and Carrel 1999 used another variable and showed an increase of 4.20 (2.18;6.22) kg of lean mass with high quality of evidence.” Please check this result, the numbers on Figure 8 are different.</p> <p>21) Results page 10: Head circumference - 0.55 (0.25; -0.86) cm compared, there is no minus sign on Figure 9.</p> <p>22) Results, page 10: Part of a figure legend seems to be present between Head Circumference and Adverse Events</p> <p>23) Excluded studies (Supplementary File 3): Six studies have been excluded due to ‘No SDS data.’ Was this is only reason to exclude? Could any of these studies have provided data for other outcomes?</p> <p>24) Figures: please make it clear in the titles of the figures whether randomised or non-randomised studies are included. Also some of the RoBs seem to be missing on the Figures. If any of the RoB criteria are not relevant for any of the non-randomised studies, please clearly state this within the Figure legends.</p>
--	--

REVIEWER	Reviewer name: Allegaert karel Institution and Country: KU Leuven, Belgium and Erasmus MC, Rotterdam, the Netherlands Competing interests: none
REVIEW RETURNED	01-Mar-2020

GENERAL COMMENTS	<p>I have read this paper with great interest, and value the paper as provided. My comments relate to some method related aspects and the context.</p> <p>Why do the authors use the word ‘global’ in the title.</p>
-------------------------	--

	<p>What's the add on benefit or relevance ?</p> <p>How has the list of outcome variables been constructed ? As a clinical researcher, not immediately involved in PW care beyond neonatal care, how 'relevant' are these outcome variables, like eg growth or length (a comment similar to the relevance of growth and length for idiopathic short stature cases). Are there any data that reflect patient centered or parent centered outcome, like eg behavioural outcome. I assume that (driven by product registration) the primary outcome variables likely do not cover these outcomes, but are there any observational data on this ?</p> <p>The same holds true for safety, like eg GH related risks for tumors in this specific populations ?</p> <p>Suggest to add in the introduction how the diagnosis of GH deficiency has been made (likely hormonal profile test, IGF-1 ? mediated).</p> <p>In the current paper, the authors consider 'pediatric life treatment' as a dichotomous variable. Does this mean that GH treatment should be done from neonatal or early infancy onwards until ? Is there a dose-effect or exposure-effect pattern ?</p> <p>I understand that the study protocol has been registered, but have the authors also considered to search clinical trial databases, like clinicaltrials.gov ? The same holds true for the (likely arbitrary) choice for 3.5 years for subcategories ? is this supported by any rationale ?</p> <p>Can the authors comment on potential biases related to exclusion criteria in the clinical studies, or how well can the current interpretations be extrapolated to the full PW population ?</p> <p>I understand that there are some reflections on these aspects in the discussion part of the paper, but still limited in my assessment. I assume that the authors suggest to aim for a GH registry with specific emphasis on the covariates of PW syndrome.</p> <p>Check line 13-14, pg : ? to be remove Intention to treat H) Sample size determination I) Outcome J) Early interruption K) Prognostic characteristics</p>
--	---

REVIEWER	<p>Reviewer name: Antonino Crinò Institution and Country: Bambino Gesù Hospital, research Institute, Rome, Italy Competing interests: Genetic obesity and in particular Prader-Willi syndrome</p>
REVIEW RETURNED	15-Mar-2020

GENERAL COMMENTS	<p>The authors performed a systematic review and, where possible, a meta-analysis on the use of growth hormone (GH) for the treatment of patients with Prader-Willi syndrome. Although a great effort has been made to carried out this interesting study, the data are not clear and are rather confusing. There are too many figures and supplementary files that do not make the reading of the text comprehensible.</p> <p>However, I have the following comments =</p>
-------------------------	---

	<p>Title: I propose to add in the title “in children and adolescents” because in this review adult patients were excluded – global? why?</p> <p>Keywords: Endocrinology should be deleted; replace growth with “growth hormone” and syndrome with “Prader-Willi syndrome”; replace also obesity with genetic obesity.</p> <p>Abstract (page 2): Line 47: The word “Method” is repeated twice RCT = please, specify the meaning of RCT Line 54: specify 1,54-1,81 etc (is it a range?)</p> <p>Introduction (Page 3): Line 19-23 = explain better the characteristic of the syndrome. In particular the ipotonia and other features of PWS are not reported Line 39 = metabolism what does it mean? Please, can you explain better</p> <p>Objective Page 4 line 7 = systematic review followed by meta-analysis (in the abstract is written “where possible”) - you need to be more homogeneous and precise.</p> <p>Methods (Page 4-6) Page 4 Line 26 = why adult studies are not included? Line 29 = please, add the abbreviations (RCTs and NRCT) Line 34 = please more information about placebo group are necessary (what does it mean?) Line 49 = the reference on DEXA is missing Line 57 = the reference on head circumference (Z-score) is missing Page 6 line 12 = the phrase is repeated</p> <p>Results (Pages 6-10): The description of data is not quite clear. How many patients are evaluated (under 3,5 yrs and older)? how many patients were with GHD? How many of them were tested? Which tests are been carried out? Page 6 Line 36 (after databases (insert Medline and Embase – no Cochrane) Page 6 Line 38: please, specify that the excluded studies were n=18 (see Supplementary file 3) Page 6 Line 41 = the RCTs are 16 and not 17 Page 6 Line 41: why only 9 of the 17 (16) RCTs have undergone the meta-analysis? Page 6 Line 47 = please, insert the reference of the studies (quality of life and muscular strength). Page 7- Line 54-55 (please insert the references of the studies); The sentences of Lines 55-58 are not clear. Page 9 - Line 4-8: the reference of the studies should be reported. – The sentences of line 7-8 are not clear. Page 10 - Line 16-26 = The description of adverse events needs to be made in more detail</p> <p>In the Supplementary 1: please, correct the pages (right side)</p> <p>In the Supplementary file 3: you can add “n:18 excluded studies after full test analysis”. How many of them were RCTs and how many NRCT studies?</p>
--	--

	<p>In the Supplementary 4: Line 4: Change cognition with “Cognitive function” (put it on the left side of the table) – also “Behaviour outcome” (Line 35) instead than Behaviour; Change motor with “motor development” (Line 44). Regarding motor n. 7 studies are reported, but in the text it is written that there are only 3 studies (please comment and correct). Please, explain the acronym of AIMS and GMFM (line 52, page 30) and BSID-II (line 54, page 30).</p> <p>Table 1 = please insert the number of patients for each study - In the title, please insert the numbers of the studies (n =16) – please, change Dose with “dose of GH”.</p> <p>Table 2 = please insert the number of patients for each study; It is not clear the difference between “before and after” and “cohort”. You must explain. In the title of Table 2 please insert the numbers of the studies (n =20); what does it mean 1:2 – 01;2 (regarding the period); please, correct (it is not clear).</p> <p>Supplementary 5: the data shown are not clear and understandable; they must be explained in the text.</p> <p>The figures (in total n=10) are too many and contain too much data. They are understandable.</p> <p>The discussion is quite short and reports what is already known in the literature. The authors should make comments on implications for future research and clinical practice.</p> <p>Being a meta-analysis the references should be expanded. The references of the excluded studies reported in Supplementary file 3 are missing.</p>
--	--

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Comments to the Author

I have performed a statistical review of the manuscript "Growth Hormone treatment in Prader-Willi syndrome patients: a global systematic review and meta-analysis"

The authors perform a systematic review and meta-analysis of randomised and non-randomised studies examining growth hormone treatment in Prader-Willi syndrome and conclude that growth hormone treatment results in an improvement of stature, body composition and BMI.

I am unclear regarding certain elements of the methods employed:

1) I do not understand exactly which study designs are eligible for this review. It is stated on page 4 that randomized and non-randomized (cohort and before and after studies) are eligible designs.

However, it is also stated in the next paragraph that rhGH is to be compared to placebo or no drugs – this comparison cannot be made in before and after studies.

Also, a quality assessment tool for case series is mentioned on page 5 – were case series also included?

Further, within the results section it is stated that “For the meta-analysis process and for a better approach to all data, we transformed all cohort studies into before and after studies.” I do not understand why or how this has been done.

Please clarify the above points regarding the study designs

Thanks for your comment.

Our objective is to include a great number of studies in PWS literature. We selected randomized and non-randomized (cohort and before and after studies) for eligible design.

For randomized studies we included articles that compared rhGH to placebo or no drugs.

For non-randomized, as you said for before and after studies there is no comparison, but for cohort we have a group that use GH and another group that used no drugs.

We write by mistake, we used cohort studies and before and after studies and no case series.

For the metanalysis we just take 1 arm of cohort studies (the group that used growth hormone: we took the data at the start and at the end of the study) transforming cohort studies in before and after studies. I made these changes in the method.

2) I do not understand most of the content under the heading “Measures of treatment effect and units of analyses, ” particularly the following sentences:

“Different follow-up lengths were standardized by use of incidence densities of outcomes occurrence across patient-years followed. When different measures or instruments were used in the studies, standard errors were used to estimate outcomes”

And also the formula to convert 25th and 75th percentiles. It seems unlikely a study would report a mean and percentiles (generally a median would be reported with percentiles).

Please clarify and provide a reference for the formula.

In before and after studies we found different follow-up periods. Due to that we standardized our measure in patient-year. It means the number of patients multiplied by years of follow-up. For example, one article has 32 patients, the mean follow-up is 8 y: the measure is $32 * 8 = 256$ patients-year.

Related to the formula, we found some articles that use the p25 and P75 as Festen 2008 and we need to do this conversion. I corrected the formula. I used median and not mean. We used the Hozo et al's method with the median. We cited the reference.

Hozo SP, Djulbegovic B, Hozo I: Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol. 2005, 5: 13-10.1186/1471-2288-5-13.

We appreciate your comment.

3) Please provide rationale for the subgroup analysis and how the cut off of 3.5 years was determined

There are many articles in literature that advice start the growth hormone during infancy.

When we collect our data, we saw that the cut-off of the majority of the randomized studies, so we use this to divide the groups.

4) The authors are correct to perform separate meta-analyses of randomised and non-randomised studies but please clearly state this within the methods.

We write in methods (measures of treatment effects): We performed separated meta-analyses of RCTs and NRCTs.

5) Please check the GRADE certainty judgements in Supplementary Files 5 and 6. For example, the analysis of Z-stature with serious risk of bias and very serious inconsistency should probably be graded lower than moderate.

Ok, I checked and. According to GRADE PRO it is moderate for all patients and high for the subgroups NRCTs and low in observational studies

How were levels of association determined as 'strong' and 'very strong'? Please provide some justification (ideally with a reference) for this.

In gradepro I informed the following data: publication bias, large effect (no, large or very large), plausible confounders and dose response gradient (that we don't have in this case) .

- Strong evidence of association—significant relative risk of > 2 (< 0.5) based on consistent evidence from two or more observational studies, with no plausible confounders (+1)
- Very strong evidence of association—significant relative risk of > 5 (< 0.2) based on direct evidence with no major threats to validity (+2)

Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328(7454):1490. doi:10.1136/bmj.328.7454.1490.

I also have some minor comments

1) I don't understand the use of the word 'global' in the title and what this refers to. I also suggest that the word 'recombinant' should be used in the title. We took off the term global

2) The terms 'somatropin' and 'rhGH' are used interchangeably. From my understanding, both terms refer to the same treatment but for consistency, I suggest choosing just one of these terms and using it throughout the manuscript

Ok, we modified.

3) Introduction: I suggest that the first paragraph of the introduction may need some more references. Does everything in this paragraph (prevalence rate, obesity, quality of life) come from reference 1 at the end of the paragraph?

Yes. We decreased the references to enter all articles analysed.

4) Methods: "All studies that had patients who received a subcutaneous application of rhGH at any dosage, ranging from 0.3 mg/m²/day to 1 mg/m²/day, were included." I don't understand this sentence – it says that 'any dosage' would be included but then quotes a range of doses? Or is '0.3 mg/m²/day to 1 mg/m²/day' the range of dosages within the included studies? If so, I suggest stating this in the results section rather than the methods section.

We changed. We maintained the sentence any dose.

5) Methods, page 4: Are any of the listed outcomes primary outcomes of this review?

We divided in primary and secondary outcomes.

6) Methods, page 4: Quality of life and behaviour 'evaluated by different instruments.' Presumably different validated instruments?

Yes. Validated instruments, we will include this.

7) Methods, page 5: "All articles PWS were evaluated independently by 2 reviewers." PWS is not needed in this sentence

Thanks, we changed. All PWS articles.

8) Methods, page 5 (Selection of studies and data extraction): Please provide details of which outcomes are analysed as dichotomous variables and which are analysed as continuous variables.

Ok, we provide the data. "Extraction of dichotomous (adverse effects) and continuous variables (all other outcomes) was independently performed."

9) Table 1 and Table 2: I'm not sure why 'systematic review 2019' is in the header of these tables?

Thanks for your comment. We took these off.

10) Table 1: Please provide details of the control groups in these studies. It would also be helpful to include the number of people in the rhGH and control groups in these tables.

We chose to put the number of patients in each Figure with all information on it.

11) Results: "We performed the meta-analysis for the evaluation of growth hormone treatment versus placebo/no treatment in nine of the seventeen included RCTs" In the previous sentence, it is stated that 16 studies were included, not 17?

I change for the correct number. 16

"We performed the meta-analysis for the evaluation of growth hormone treatment versus placebo/no treatment in nine of the sixteen included RCTs"

Also, why could only 9 studies be included? Were appropriate data not available in the other studies?

It was related to the outcome.

Some studies don't have the same outcome to evaluate. Due to that we made a table with the outcomes that could not be evaluated. I explain this in results "The seven meta-analysis excluded studies contain outcomes related to cognitive function, behaviour, quality of life and motor development."

12) Results: "For the evaluation of cognitive function in patients with PWS, different questionnaires were used; therefore, global analysis was not possible." I'm not sure what a 'global' analysis is, should this read 'meta-analysis'?

We changed for meta-analysis.

13) Assessment of heterogeneity and reporting bias, page 6: I-squared is a statistic rather than a test

Ok, we changed

14) Assessment of heterogeneity and reporting bias, page 6: Is the p-value in this section relating to a test of the heterogeneity or of the treatment effect?

It is related to treatment effect. We changed for prior section.

15) Assessment of heterogeneity and reporting bias, page 6: "The possibility of publication bias was assessed by evaluating a funnel plot (Begger Test) for asymmetry" I cannot find a funnel plot within the Figures? Was there any evidence of publication bias?

We didn't include the figure but we made a funnel plot, and I can send it if you want. There was no publication bias.

16) Results, page 9: When quoting results in the text please clearly state that the numbers in brackets are a confidence interval, for example 1.67 (95% CI 1.54-1.81) SDS. Please also clearly state in the text where heterogeneity is present.

Thanks, we changed in all the text.

17) Results, page 9: "In the NRCT studies, the treated group had a significant improvement of stature (1.52 (0.86;2.16) SDS) during the 2,909 patients-years of follow-up (Figure 2)." Please check this result, the numbers on Figure 2 are different.

The number is correct. The minus sign is because stature starts negative. To show the difference we use a positive result and not a negative.

18) Results, page 9: "We chose to report the growth velocity measure in centimeters and due to that one article that used SDS was excluded." I do not understand what this means?

It means that we excluded one article due to a different measure centimeters and centimeters in SDS.

We changed here to be clearly "We chose to report the growth velocity measure in centimeters and due to that one article that used only SDS measure was excluded."

19) Results, page 9: Lean mass – two studies seem to be included in each analysis but the text states three.

It was a mistake. I made the correction, two is correct.

20) Results, page 9-10: "The studies by Lindgreen 1998, and Carrel 1999 used another variable and showed an increase of 4.20 (2.18;6.22) kg of lean mass with high quality of evidence." Please check this result, the numbers on Figure 8 are different.

Ok, we changed. Thanks

21) Results page 10: Head circumference - 0.55 (0.25; -0.86) cm compared, there is no minus sign on Figure 9.

Ok, we changed the minus sign. It is positive

22) Results, page 10: Part of a figure legend seems to be present between Head Circumference and Adverse Events

Ok, we changed

23) Excluded studies (Supplementary File 3): Six studies have been excluded due to 'No SDS data.' Was this is only reason to exclude? Could any of these studies have provided data for other outcomes?

Unfortunately, these studies just contains the means of all outcomes. No other outcomes could be used.

24) Figures: please make it clear in the titles of the figures whether randomised or non-randomised studies are included. Also some of the RoBs seem to be missing on the Figures. If any of the RoB criteria are not relevant for any of the non-randomised studies, please clearly state this within the Figure legends.

Ok, I correct the figures titles.

When is missing if it is red or green that means that we cannot made a conclusion regarding the item. For example in the case o RCTs. There are some that use envelopes for randomization, it is not red, and is not green because is not computerized, so it is yellow but we don't have this option in Revman, so we put that it is not clear.

Reviewer: 2

Comments to the Author

I have read this paper with great interest, and value the paper as provided. My comments relate to some method related aspects and the context.

Why do the authors use the word' global' in the title. What's the add on benefit or relevance ?

Change global- included

Combined?? Recombinant?

We suppressed the word global.

How has the list of outcome variables been constructed ? As a clinical researcher, not immediately involved in PW care beyond neonatal care, how 'relevant' are these outcome variables, like eg growth or length (a comment similar to the relevance of growth and length for idiopathic short stature cases).

The list of outcomes was based on clinical care of PWS patients. They are the most important outcomes for these patients. Growth is important is these cases because most of them have growth hormone deficiency so they can be very short. But is not the main outcome. Due to that we select other outcomes.

Are there any data that reflect patient centered or parent centered outcome, like eg behavioural outcome. I assume that (driven by product registration) the primary outcome variables likely do not cover these outcomes, but are there any observational data on this ?

The same holds true for safety, like eg GH related risks for tumors in this specific populations ?

Because we want to discuss multiple outcomes, we included cognition, motor development , behavior and quality of life. Regarding bevahior outcomes, but we didn't find so many differences. Some studies comment that with GH children and adults are less tired and want to do more physical activity with rhGH. Also, cognitive outcome is centered in the patient. We see some good results in these field. I expanded my explanation about this in the discussion part.

Gh is related to tumors in patients that have a previous cancer. They have a great chance to have another. In the other population GH is not related with cancer.

Suggest to add in the introduction how the diagnosis of GH deficiency has been made (likely hormonal profile test, IGF-1 ? mediated).

I included the GH stimulated test in the sentence, but the diagnosis of GH deficiency is very controversy. Growth hormone (GH) deficiency can occur in 40 to 100% of PWS patients, depending on GH stimulated test. But in fact our main focus is that “The use of rhGH is currently recommended in PWS patients with and without GH deficiency to improve short stature and metabolism”

In the current paper, the authors consider ‘pediatric life treatment’ as a dichotomous variable. Does this mean that GH treatment should be done from neonatal or early infancy onwards until ? Is there a dose-effect or exposure-effect pattern?

Yes, it could be started in early infancy at 2 or 3 months of age. But generally, before they are hospitalized and is more complicate to introduce the medication and there are some priorities. But if the child is presenting hypoglycemia it will be great to introduce earlier.

Dose effect was not possible to evaluate in our trial. Most of the centers did the same dose.

I understand that the study protocol has been registered, but have the authors also considered to search clinical trial databases, like clinicaltrials.gov ? The same holds true for the (likely arbitrary) choice for 3.5 years for subcategories ? is this supported by any rationale ?

We look for these studies in clinical trials , but unfortunately most of the countries that do research in WS are not authorized to do Randomized studies due ethical reasons, because is approved, but there are other many countrues that are not approved.

Can the authors comment on potential biases related to exclusion criteria in the clinical studies, or how well can the current interpretations be extrapolated to the full PW population ?

Ok, we included this comment in the discussion part.

I understand that there are some reflections on these aspects in the discussion part of the paper, but still limited in my assessment. I assume that the authors suggest to aim for a GH registry with specific emphasis on the covariates of PW syndrome.

Thank you, we included your assessment.

Check line 13-14, pg : ? to be remove

Intention to treat H) Sample size determination I) Outcome J) Early interruption K) Prognostic characteristics

Ok, we removed

Reviewer: 3

Comments to the Author

The authors performed a systematic review and, where possible, a meta-analysis on the use of growth hormone (GH) for the treatment of patients with Prader-Willi syndrome. Although a great effort has been made to carried out this interesting study, the data are not clear and are rather confusing. There are too many figures and supplementary files that do not make the reading of the text comprehensible.

However, I have the following comments =

Title:

I propose to add in the title “in children and adolescents” because in this review adult patients were excluded – global? why?

Ok, accepted.

Adults are not included because there are few studies with different outcomes, and we have already a lot of data.

Keywords:

Endocrinology should be deleted OK; replace growth with “growth hormone” and syndrome with “Prader-Willi syndrome”; replace also obesity with genetic obesity.

Abstract (page 2):

Line 47: The word “Method” is repeated twice

It was a mistake, we changed

RCT = please, specify the meaning of RCT

Ok, we changed

Line 54: specify 1,54-1,81 etc (is it a range?)

We changed to 95% CI- confidential interval

Introduction (Page 3):

Line 19-23 = explain better the characteristic of the syndrome. In particular the ipotonia and other features of PWS are not reported

We chose to explain the hypotonia later because is not the main focus of our study.

Line 39 = metabolism what does it mean? Please, can you explain better

In the next phrase in the text I explain that “ GH acts as an anabolic agent that increases lean body mass (LBM) and reduces fat mass. “ what explain their action in metabolism.

Objective

Page 4 line 7 = systematic review followed by meta-analysis (in the abstract is written “where possible”) - you need to be more homogeneous and precise.

Ok, we changed

Methods (Page 4-6)

Page 4 Line 26 = why adult studies are not included?

Our main focus were children and we have already a lot of data and outcomes to evaluate. There are few studies in adults to do a metanalysis.

Line 29 = please, add the abbreviations (RCTs and NRCT). We added

Line 34 = please more information about placebo group are necessary (what does it mean?).

It means no drugs. No study used placebo.

Line 49 = the reference on DEXA is missing.

We included (DXA, type Lunar Prodigy, GE Healthcare, Chalfont St Giles, UK) (DEXA; Lunar Prodigy by GE Medical System, Madison, WI)

Line 57 = the reference on head circumference (Z-score) is missing. We included

CDC Growth Charts for the United States and Netherland charts. WE added this data.

Kuczmarski, R. & Ogden, Cynthia & Guo, S.. (2000). 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat.* 246. 1-201.

Fredriks, A., van Buuren, S. & Burgmeijer, R. (2000) Continuing positive secular growth change in the Netherlands 1955 –97. *Pediatric Research*, 47, 316–323.

Page 6 line 12 = the phrase is repeated. Ok

Results (Pages 6-10):

The description of data is not quite clear. How many patients are evaluated (under 3,5 yrs and older)? how many patients were with GHD? How many of them were tested? Which tests are been carried out?

We included all patients during childhood and adolescence despite of the GH stimulatory test. We included the number of patients in the table.

Most of the articles don't divided patients by the test, because there are different types and they are very controversial in literature, and they depend on age. So, we chose the same option them the articles that the majority included articles.

Page 6 Line 36 (after databases (insert Medline and Embase – no Cochrane)

Page 6 Line 38: please, specify that the excluded studies were n=18 (see Supplementary file 3). We included.

Page 6 Line 41 = the RCTs are 16 and not 17. We changed

Page 6 Line 41: why only 9 of the 17 (16) RCTs have undergone the meta-analysis?

It was related to the outcome.

Some studies don't have the same outcome to evaluate. Due to that we made a table with the outcomes that could not be evaluated. We explained better this part.

“The seven metanalysis excluded studies contain outcomes related to cognitive function, behavior, quality of life and motor development.”

Page 6 Line 47 = please, insert the reference of the studies (quality of life and muscular strength).

We included

Page 7- Line 54-55 (please insert the references of the studies); The sentences of Lines 55-58 are not clear.

We included

Page 9 - Line 4-8: the reference of the studies should be reported. – The sentences of line 7-8 are not clear.

We included

Page 10 - Line 16-26 = The description of adverse events needs to be made in more detail.

Ok, we included.

In the Supplementary 1: please, correct the pages (right side)

In the Supplementary file 3: you can add "n:18 excluded studies after full test analysis". How many of them were RCTs and how many NRCT studies? Ok, we included. There is just 4 RCTs: Myers 1999 (same as Carrel 2000); Festen 2007 (same data as 2008); Haqq and Whitman that are a 6 months trial.

All others are NRCTs.

In the Supplementary 4: Line 4: Change cognition with "Cognitive function" (put it on the left side of the table) – also "Behaviour outcome" (Line 35) instead than Behaviour; Change motor with "motor development" (Line 44). Regarding motor n. 7 studies are reported, but in the text it is written that there are only 3 studies (please comment and correct). 3 RCTs studies and 4 NRCTs

Please, explain the acronym of AIMS and GMFM (line 52, page 30) and BSID-II (line 54, page 30).

They are in the table at the right side, but I made this clearly. Muscle strength (Infant Muscle Strength Meter- AIMS) and Motor function (Gross Motor Function Measure-GMFM)

Table 1 = please insert the number of patients for each study - In the title, please insert the numbers of the studies (n =16) – please, change Dose with "dose of GH".Ok, we corrected- dose of GH. How the number of patients change in the same study regarding the outcome. We chose to put in the figures.

Table 2 = please insert the number of patients for each study; It is not clear the difference between "before and after" and "cohort". You must explain. In the title of Table 2 please insert the numbers of the studies (n =20); what does it mean 1:2 – 01;2 (regarding the period); please, correct (it is not clear).

It is regarding study period. It is in the upper part of the table (1;2 years). The number of patients is now included in the table.

Supplementary 5: the data shown are not clear and understandable; they must be explained in the text.

The data showed came from Grade website that is the main table for evidence in metanalysis.

The figures (in total n=10) are too many and contain too much data. They are understandable.

Unfortunately, many figures are necessary. They provide all the data that is necessary for this metanalysis.

The discussion is quite short and reports what is already known in the literature. The authors should make comments on implications for future research and clinical practice.

Ok, we included

Being a meta-analysis the references should be expanded. The references of the excluded studies reported in Supplementary file 3 are missing.

Include excluded articles

We include references for excluded articles.