PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Cognitive and motor skills in school-aged children following maternal vitamin A supplementation during pregnancy in rural Nepal: A follow-up of a placebo-controlled, randomized cohort.
AUTHORS	Buckley, Gillian; Murray-Kolb, Laura; LeClerq, Steven; Khatry, Subarna; West, Keith; Wu, Lee; Christian, Parul

VERSION 1 - REVIEW

REVIEWER	Lora Iannotti, Ph.D. Assistant Professor, Institute for Public Health George Warren Brown School of Social Work Washington University in St. Louis USA I have no conflicts of interest in this work.
REVIEW RETURNED	03-Jan-2013

GENERAL COMMENTS	This study is a fascinating look at whether maternal vitamin A supplementation effects long-term cognitive and motor development outcomes in the offspring. Investigators have tackled a challenging question and should be congratulated on the effort. Below, please find some general comments regarding overall concepts and the study design, and then more specific comments with areas for improvement within the manuscript. General Comments Study design and methods. To begin, it would be helpful if more explanation was provided for why this study was originally designed to be cluster-randomized since a placebo was used. Relatedly, authors aptly apply statistical techniques to handle the cluster design, but I was unsatisfied with the accounting for confounding factors that might arise from clustering at the VDC/ward levels. Even if these factors were not collected in the original trial or the follow-up, acknowledgement of their possible role would be important in the discussion section for example. One important factor that is not covered is variability in schools and teacher quality. It was unclear how many different schools might be present in the VDC, or whether there are private or public differences. It seems unlikely that standardized tests/scores would be used in this context, but this sort of variable might have been useful. As well, the micro- and macro-economic conditions potentially influencing diet quality and/or schooling were not fully considered. Authors mentioned land-ownership differences in the vitamin A group participation, but it is unclear if the variable was considered in the adjustment for effects? Finally, seasonality is included as a co-factor, but the year of supplementation might also

be examined in modeling effects.

Multiple nutrient deficiencies and brain development pathways. Another general comment would be that the manuscript neglects the possibility for vitamin A deficiency in this population to be occurring concurrently with other micronutrient deficiencies that may be affecting brain development. For example, iron and zinc in particular are critical for early brain development in overlapping areas expressed for vitamin A in the manuscript (hippocampus, learning and memory) and probably missing in the diets of these children as evidenced by the very high prevalence of stunting for example. Fatty acids are also vital in early brain development but not mentioned in the paper. The effects of vitamin A supplementation, thus, could be attenuated by deficiencies in these other nutrients -- especially if there are shared pathways.

Related to this issue, is what authors describe in on line 46 of page 7 in the Methods section only about VAS program in Nepal beginning in 1995. It seems likely that the majority of these children could have received VAS during childhood. In view of the wide confidence intervals for cognitive and motor development, and potential for brain plasticity, this could also have attenuated the effects. I think it deserves more attention in the Discussion section.

UNIT & MABC tests. The final general comment would be that authors need to discuss further the literature for use of these tests in detecting cognitive and motor development effects. More references are needed throughout the document on its application (or text describing its lack of application). For example, it is unclear how the “40% of SD” hypothesized effect was determined. Also while authors do a nice job discussing the possibility for imprecision with the tests in the Discussion section (p. 16, line 39-40), no mention is made for how the effect estimate and resulting sample size may have influenced the lack of significance in the findings.

Specific comments:

p. 4, line 18 Please specify what “This” is referring to in this bullet point

p. 5, lines 18-20 Excellent sentence! The paper is well-written.

p. 5, line 44 Do you mean “chronic dietary vitamin A deficiency” instead of “deficit”?

p. 6, line 30 “cardiovascular”

p. 6, lines 46-56 These sentences are somewhat confusing. Please clarify the unit of randomization, the ward?

p. 7, line 20 I was not entirely convinced by the rationale for not including the beta-carotene arm. Maybe also due to logistics/cost?

p. 7, lines 42-56 Again, VAS program could have attenuated the effects and should be mentioned again in the Discussion section (see general comment).

p. 8, line 20 What do you mean by “fair”? Should try to find a more objective descriptor perhaps.

	p. 9, lines 17-25 Need reference or more rationale for effect size estimate. p. 10, line 3 Again, terms are somewhat subjective, “comfortable” and “simple”. The sentence may not be needed? p. 12, line 37 Not clear why missing history of night blindness excluded children since trial was randomized. p. 13, line 8 Shouldn’t land ownership be adjusted for in final models? (see general comments) p. 13, line 48-51 Anemia prevalence reporting is different in the table compared to the text?? p. 14, line 15. Perhaps use “stunted” instead of “short” p. 14, line 56 Why season but not year? Likely to be differences in food availability/access, economic and political conditions, healthcare access, etc. (see general comment) p. 16, line 17-30 Perhaps discuss role of other nutrients acting together with vitamin A p. 16, lines 36-55 Include here the possibility for being under-powered (see general comment). p. 17 Could include text in the Discussion describing other confounding – school/teacher quality, SES, maturation effects – and possibly the role of the VAS supplementation program in Nepal attenuating effects due to brain plasticity.

REVIEWER	GV Krishnaveni Research Scientist and Wellcome Trust Intermediate Fellow Epidemiology Research Unit CSI Holdsworth Memorial Hospital Mysore, India
REVIEW RETURNED	04-Jan-2013

THE STUDY	As I am do not have enough expertise to assess the statistical tests used, I suggest a statistical review.
GENERAL COMMENTS	Reviewer’s comments: The paper by Buckley et al. reports the association between maternal peri-conceptional vitamin A supplement status, and offspring cognitive performance in later childhood. This is one of the few studies examining the long-term causal effects of maternal nutrition, particularly micronutrients, on offspring outcomes. Thus, this adds to a very limited literature in this field. General comments: 1. Overall, the language used is complex and makes a difficult reading at several places for those whose first language is not English. The sentences are long, and in many places the citation numbers appear in the middle, which break the flow of the reading. I

	suggest simplifying the text, and keep the sentences short and succinct. Specific comments: 1. In the Results section it has been given that 751 babies were born in the study wards selected for this follow-up. Please start from this number in your flow chart; this helps better understanding of the sample selection procedure. Also, only 485 (64%) of the 751 babies were available for randomization in the end. Please discuss whether this has created any bias and influenced your findings? 2. Information on whether the cognitive tests have been adapted and validated for the local culture, and if yes how was this done is needed in the Methods section. 3. What were the components of the SES measured? Whether parental education, especially maternal education, which is an important predictor of children's cognitive functions, has been measured? 4. Please include description on Raven's Coloured Progressive Matrices; what does this measure? 5. The statistical methods section should include the calculation of Z-scores. 6. It will be useful to discuss the factors that might have influenced their findings in the Discussion section. Eg: is it due to residual confounding such as maternal education, children's current nutrient status? As the authors state, all these children may have been exposed to vitamin A supplementation during early childhood as part of a national programme, which might have had a far reaching effect on their cognitive development. 7. What did the placebo contain? It is important to include a sentence on this in the manuscript though it has been described in the original study. 8. The women were supplemented even before pregnancy. How long before pregnancy was this and was the time of pre-conceptional supplementation uniform for all? 9. Tables 3 and 4: Please add foot notes to describe how the effect size was estimated and what does the p-value stand for.

REVIEWER	Fowzia Ibrahim King's Musculoskeletal Clinical Trials Unit Department of Academic Rheumatology King's College London Weston Education Centre Cutcombe Road London SE5 9RJ
REVIEW RETURNED	14-Feb-2013

GENERAL COMMENTS	Page 11- The potential confounder factors that was adjusted is in the results section, I would suggest moving to statistical analysis section. Presumably SAS was used to analyse the data could this be added to statistical analysis section as well. Also could the authors specify that they have used 2-sided p-values throughout the results? It will also be helpful to mention that there was no missing data on the outcome 3 subjects all together had no MABC)- quite an achievement. However, there was quite few percentage of missing explanatory variables, which had some implication in the multivariate

	analysis. There ought to be some mention on this in the results or discussion section. Page-12: I would suggest a small heading after line 46, “Cohort or participant description” since this paragraph describes the combine sample of 381 rather than the individual groups. Page-14: It will be helpful to transform back the MABC to its original scale when reporting in the results but either way its fine statistically, just wondering how noon-statistician will interpret it. Table 2- some of the questions i.e. contemporary SES, presumably the subjects can choose more than one question or are those variable separate, for example, Has Latrine (Yes: 47 (23%) and thus 77% are No). This is not obvious in tables, could the authors make more apparent by putting a foot note or expanding the table to have more rows that show Yes/No. Well done to authors, the articles read very well and very informative, specially reporting the missing data in the explanatory variables.

VERSION 1 – AUTHOR RESPONSE

Reviewer: Lora Iannotti, Ph.D.
Assistant Professor, Institute for Public Health
George Warren Brown School of Social Work
Washington University in St. Louis
USA

I have no conflicts of interest in this work.

Please see general comments regarding consideration of other confounding factors such as school & teacher variability, year of supplementation, and economic factors. It may not be possible to adjust for these factors if not measured, but could be acknowledged in the discussion section.

Please see general comments about the interpretation of results with a view towards multiple micronutrient deficiencies, the UNIT/MABC literature, and the VAS program in Nepal.

This study is a fascinating look at whether maternal vitamin A supplementation effects long-term cognitive and motor development outcomes in the offspring. Investigators have tackled a challenging question and should be congratulated on the effort. Below, please find some general comments regarding overall concepts and the study design, and then more specific comments with areas for improvement within the manuscript.

General Comments

Study design and methods. To begin, it would be helpful if more explanation was provided for why this study was originally designed to be cluster-randomized since a placebo was used. Relatedly, authors aptly apply statistical techniques to handle the cluster design, but I was unsatisfied with the accounting for confounding factors that might arise from clustering at the VDC/ward levels. Even if these factors were not collected in the original trial or the follow-up, acknowledgement of their possible role would be important in the discussion section for example.
A better explanation for the cluster-randomized design is added to page 5. The researchers used a

placebo control design because using individual randomization for so large a sample would introduce the chance of contamination. (That is, the chance that women may receive a treatment other than the one they had been assigned to.) The sample size for the original trial was also large enough to make the cluster-randomized design efficient; the NNIPS-2 trial enrolled ~45,000 women from 270 wards.
In response to the concern with clustering of certain factors at the ward level, we added to the manuscript a mention of the possibility that other factors difficult to measure might have influenced the results. The most likely hidden variable is quality of schooling. Although we measured the child's school history, the quality of the school is difficult to measure in Nepal, where there are no standardized ways to evaluate schools. Given that we did measure the quality of the environmental stimulation, the child's school history, mother's intellectual ability, and various other social and nutritional influences on development, and found no differences between groups, it seems implausible that school quality would be so imbalanced between groups.

One important factor that is not covered is variability in schools and teacher quality. It was unclear how many different schools might be present in the VDC, or whether there are private or public differences. It seems unlikely that standardized tests/scores would be used in this context, but this sort of variable might have been useful. As well, the micro- and macro-economic conditions potentially influencing diet quality and/or schooling were not fully considered. Authors mentioned land-ownership differences in the vitamin A group participation, but it is unclear if the variable was considered in the adjustment for effects? Finally, seasonality is included as a co-factor, but the year of supplementation might also be examined in modeling effects.
As noted above, there are no standardized test scores in Nepal, though we agree that such a measure would be useful in distinguishing between better and poorer quality schools. The land ownership variable was not adjusted for in the adjusted analysis because other measures of socio-economic status measured in 2008 were used instead. To have controlled for land ownership as well would have amounted to over-controlling for the effects of the same variable (socio-economic status) in the analysis.
We also believe it would be unwise to adjust for year. Data collection took only nine months. Children tested later were generally older at testing then those tested first. The adjusted models did control for age, and age would be collinear with year if that variable would included in the model.

Multiple nutrient deficiencies and brain development pathways. Another general comment would be that the manuscript neglects the possibility for vitamin A deficiency in this population to be occurring concurrently with other micronutrient deficiencies that may be affecting brain development. For example, iron and zinc in particular are critical for early brain development in overlapping areas expressed for vitamin A in the manuscript (hippocampus, learning and memory) and probably missing in the diets of these children as evidenced by the very high prevalence of stunting for example. Fatty acids are also vital in early brain development but not mentioned in the paper. The effects of vitamin A supplementation, thus, could be attenuated by deficiencies in these other nutrients -- especially if there are shared pathways.

In response to this comment the discussion section is expanded to include a mention of how other nutritional deficiencies might have influenced the results. See page 16 and response to specific comment below.

Related to this issue, is what authors describe in on line 46 of page 7 in the Methods section only about VAS program in Nepal beginning in 1995. It seems likely that the majority of these children could have received VAS during childhood. In view of the wide confidence intervals for cognitive and motor development, and potential for brain plasticity, this could also have attenuated the effects. I think it deserves more attention in the Discussion section.
In response to this comment the role of the national vitamin A supplementation program is described in better detail on page 16. This addition explains that the supplementation program might have

attenuated a treatment effect, allowing placebo group children to regain the nutritional losses of their early life.

UNIT & MABC tests. The final general comment would be that authors need to discuss further the literature for use of these tests in detecting cognitive and motor development effects. More references are needed throughout the document on its application (or text describing its lack of application). For example, it is unclear how the “40% of SD” hypothesized effect was determined. Also while authors do a nice job discussing the possibility for imprecision with the tests in the Discussion section (p. 16, line 39-40), no mention is made for how the effect estimate and resulting sample size may have influenced the lack of significance in the findings.

In response to this comment we describe better how the UNIT is a relatively new psychometric test, and this study is the first known use of the UNIT in Nepal. The MABC is an older psychometric instrument, but we are unaware of other research groups using the test in Nepal.
While the tests were not formally validated before this project, a team of school psychologists from Penn State University closely monitored the process validity of the testing. And, as explained in a revision on page 8, this study aimed only to compare groups, not to diagnose disability. Therefore, the lack of a normative sample from which to calculate normalized scores is not a serious limitation.
Regarding the reviewer’s comment on the sample size, we powered the study to see a difference of about 40% of a standard deviation. There is no normative data on the UNIT and MABC tests from Nepal, we could not calculate standardized scores and thought it made the most sense to express the expected change as a percentage of a standard deviation. A change of one third of a standard deviation is generally regarded as meaningful in the clinical literature. We considered the detectable difference of 40% meaningful.

Specific comments:

p. 4, line 18 Please specify what “This” is referring to in this bullet point
The bullet point is revised. (Please note this box appears at the beginning of the reviewer copy, but at the end of the submission and resubmission.)

p. 5, lines 18-20 Excellent sentence! The paper is well-written.
Thank you

p. 5, line 44 Do you mean “chronic dietary vitamin A deficiency” instead of “deficit”?
Yes, deficit is revised to deficiency.

p. 6, line 30 “cardiovascular”
Revised.
p. 6, lines 46-56 These sentences are somewhat confusing. Please clarify the unit of randomization, the ward?
The paragraph was revised; unit of randomization is now mentioned explicitly. See page 6, highlighted paragraphs.

p. 7, line 20 I was not entirely convinced by the rationale for not including the beta-carotene arm. Maybe also due to logistics/cost?
The reviewer is correct that cost and logistical constraints factored into the decision, mention of these considerations is added to page 6, the second paragraph.

p. 7, lines 42-56 Again, VAS program could have attenuated the effects and should be mentioned again in the Discussion section (see general comment).

At the reviewers suggestion a better discussion of the supplementation program is added to the Discussion section, see page 16 second paragraph.

p. 8, line 20 What do you mean by “fair”? Should try to find a more objective descriptor perhaps.
This descriptor is revised to “accessible”.
p. 9, lines 17-25 Need reference or more rationale for effect size estimate.
In response to this concern about the sample size, we admit in the results section that the effect size chosen was somewhat arbitrary. For better clarity, this draft contains frank mention to the cost and logistical constraints that influenced the study.

p. 10, line 3 Again, terms are somewhat subjective, “comfortable” and “simple”. The sentence may not be needed?
The sentence is not needed and removed.

p. 12, line 37 Not clear why missing history of night blindness excluded children since trial was randomized.
Night blindness was an exposure of interest and the subject of another research question in the study. A revision clarifying this point is added on page 12.

p. 13, line 8 Shouldn’t land ownership be adjusted for in final models? (see general comments)
The final models adjust for the differences in socio-economic status between groups. Table 2 and page 13 describe how the placebo group children are relatively less well-off, evidenced by a smaller proportion of them coming from households with an improved water source. Water source is the measure of socio-economic status imbalanced in the 2008 data, and this was controlled for in the analysis. To control for the land ownership variable from the 1994-1997 data as well would likely over-control for the effects of socio-economic status.
After controlling for socio-economic status, test season, age, anemia, salt iodine, there was no difference in test scores between groups.

p. 13, line 48-51 Anemia prevalence reporting is different in the table compared to the text??
The text is corrected on page 13; the order of the percentages was backwards in the earlier draft.

p. 14, line 15. Perhaps use “stunted” instead of “short”
Revised at reviewer’s suggestion.

p. 14, line 56 Why season but not year? Likely to be differences in food availability/access, economic and political conditions, healthcare access, etc. (see general comment)
Given that testing took only nine months, we consider it sufficient to consider season in the analysis, but no other calendar variable. Furthermore, year of testing would covary with the participant’s age, as the sample of children aged over the time it took to conduct the study. To control for year would cause the analysis to over-control for age.

p. 16, line 17-30 Perhaps discuss role of other nutrients acting together with vitamin A
At the reviewer’s suggestion this discussion is added, in as much detail as space will allow, to the Discussion section on page 16. Iron and zinc deficiencies are, as the reviewer mentions, common in Nepal, as is widespread malnutrition, as evidenced by essentially across-the-board stunting in this sample. Participants were clearly exposed to a wide variety of nutritional deficiencies, and the reviewer is right to point out that some of these other problems might have obscured the relationship between vitamin A deficiency and the outcomes of interest. The same section of the discussion section (pages 15-16) also brings out the importance of other nutrients to child, and the possibility that the effect of vitamin A alone may not be measurable when the deficiency occurs with other important dietary deficiencies.

p. 16, lines 36-55 Include here the possibility for being under-powered (see general comment).
A mention of this point is added to the discussion section on page 16. As the study only drew participants from 12 clusters, there is the possibility that a smaller difference in average scores might have been undetected.

p. 17 Could include text in the Discussion describing other confounding – school/teacher quality, SES, maturation effects – and possibly the role of the VAS supplementation program in Nepal attenuating effects due to brain plasticity.
The reviewer observes that there may be other factor associated with cluster, and therefore covariate with the exposure, that influence test scores. Some of these factors are difficult to measure, a point about differences in school and teacher quality is added to the discussion on pages 16-17. The two groups do not differ on the socio economic indicators measured in either the first NNIPS study or the follow-up study described. (See Tables 1 & 2.) The only variable that was different between the groups was land ownership in the 1994-1997 data. The reviewer noted this difference earlier. In response (as noted above) we note that an adjusted analysis controlled for differences in socio-economic status. These differences, presented in Table 2, and described on pages 13 and 14, are that children in the placebo group come from slightly less well-off homes; they are less likely to have an improved water source or properly iodized salt. An analysis controlling for these differences, and other imbalances such as test season, is presented in Table 2 and discussed on page 1, third paragraph. We believe that controlling for land ownership, a variable only collected in the 1994-1997 study would over control for the effects of socio-economic status.

Reviewer: GV Krishnaveni
Research Scientist and Wellcome Trust Intermediate Fellow
Epidemiology Research Unit
CSI Holdsworth Memorial Hospital
Mysore, India

As I am do not have enough expertise to assess the statistical tests used, I suggest a statistical review.

Cognitive and motor skills in school-aged children following maternal vitamin A supplementation during pregnancy in rural Nepal: A follow-up of a placebo-controlled, randomized cohort.

Reviewer's comments:

The paper by Buckley et al. reports the association between maternal peri-conceptional vitamin A supplement status, and offspring cognitive performance in later childhood. This is one of the few studies examining the long-term causal effects of maternal nutrition, particularly micronutrients, on offspring outcomes. Thus, this adds to a very limited literature in this field.

General comments:

1. Overall, the language used is complex and makes a difficult reading at several places for those whose first language is not English. The sentences are long, and in many places the citation numbers appear in the middle, which break the flow of the reading. I suggest simplifying the text, and keep the sentences short and succinct.
In response to this comment several sections are edited, especially parts of the study background and methods. These edits will improve the clarity of the paper for all readers, especially non-native English

speakers.

Specific comments:

1. In the Results section it has been given that 751 babies were born in the study wards selected for this follow-up. Please start from this number in your flow chart; this helps better understanding of the sample selection procedure. Also, only 485 (64%) of the 751 babies were available for randomization in the end. Please discuss whether this has created any bias and influenced your findings?
For better clarity, an explanation is added to the text call out for Figure 1 explaining that the flow chart shows the children alive and accounted for in 2007. In response to the point about survival bias, a mention of our comparisons between children who were and were not in the study sampling frame is added. Briefly, children in the sampling frame were less likely to have missing data in their 1994-1997 study records (missing data in 9.3% of the sampling frame, 26.6% of the early deaths, and 31.7% of those excluded from the sampling frame for being lost to follow-up or having missing data in the mother's night blindness record). As the manuscript notes on page 12, the vitamin A children in the sampling frame also came from slightly more affluent families than the placebo children.
The reviewer is correct that these differences introduce the possibility of survival bias, and for this reason, the adjusted analysis presented on page 14 and Table 3 adjust for socio-economic status. Also, for better clarity, a mention of survival bias is added to pages 16-17 .

2. Information on whether the cognitive tests have been adapted and validated for the local culture, and if yes how was this done is needed in the Methods section.
The reviewer's concern with test validity is acknowledged in a revisions on pages 7 and 8. Given that this study aimed only to compare scores between two groups of children, not to diagnose disability or calculate an IQ score, we do not see this as a serious limitation.

3. What were the components of the SES measured? Whether parental education, especially maternal education, which is an important predictor of children's cognitive functions, has been measured?
The components of socio-economic status measured are reported in Table 1 and 2. We also collected data on parent's education, however, as Table 1 shows, female illiteracy is overwhelmingly common in this population. Given that only ~10% of the mothers in this study could even read, we consider adjusting for the small additional variability in education among that minority of women unwise over-adjustment. Instead, we used the Raven's Coloured Matrices score to estimate cognitive ability in the mothers. As Table 2 shows, we found no difference in Raven's Coloured Matrices scores between groups.

4. Please include description on Raven's Coloured Progressive Matrices; what does this measure?
In response to the comment, a description of the test, one of the oldest and most validated cognitive assessments in use is added to page 9.

5. The statistical methods section should include the calculation of Z-scores.
Added at the reviewer's suggestion to page 10. Child height-for-age and BMI-for-age z-scores were calculated using the WHO NCHS 2006 growth reference.

6. It will be useful to discuss the factors that might have influenced their findings in the Discussion section. Eg: is it due to residual confounding such as maternal education, children's current nutrient status? As the authors state, all these children may have been exposed to vitamin A supplementation during early childhood as part of a national programme, which might have had a far reaching effect on their cognitive development.
In response to this comment, more detail about the routine vitamin A supplementation program is added to the Discussion on page 16.

7. What did the placebo contain? It is important to include a sentence on this in the manuscript though it has been described in the original study.
The composition of the placebo (5 mg dl-α-tocopherol and gelatin) is in the manuscript, on page 6.

8. The women were supplemented even before pregnancy. How long before pregnancy was this and was the time of pre-conceptional supplementation uniform for all?
Women were supplemented for different lengths of time, because they became pregnant at different times over the three year course of the study. The mothers of children born in 1997, for example, had been receiving supplements for years before the child was born, while mothers of children born in 1994 might have been taking it for only a few months. The length of time the woman had been supplemented before pregnancy is completely colinear with the child's age. The adjusted analysis shown in Table 3 and described on page 14 controls for age.

9. Tables 3 and 4: Please add foot notes to describe how the effect size was estimated and what does the p-value stand for.
The p-value shown in these tables is the probability of finding the result shown or one more extreme under the null hypothesis of no difference between groups. The effect estimate is the score difference between groups. A better explanation of these findings in found in the results section, on pages 14-15.

Reviewer: Fowzia Ibrahim
King's Musculoskeletal Clinical Trials Unit
Department of Academic Rheumatology
King's College London
Weston Education Centre
Cutcombe Road
London
SE5 9RJ

Page 11- The potential confounder factors that was adjusted is in the results section, I would suggest moving to statistical analysis section. Presumably SAS was used to analyse the data could this be added to statistical analysis section as well. Also could the authors specify that they have used 2-sided p-values throughout the results?
We consider it unwise to include the results of the between group comparisons in the Methods section. However, to be more clear to readers, we have changed the "Statistical Analysis" heading to "Analysis" to be more clear that it is a part of the Methods section.
Stata 9.0 (College Station, Texas) was used for all analysis. A mention of this is added to page 11.
We also added specificity to page 10 that two-sided p-values were used throughout.

It will also be helpful to mention that there was no missing data on the outcome 3 subjects all together had no MABC)- quite an achievement. However, there was quite few percentage of missing explanatory variables, which had some implication in the multivariate analysis. There ought to be some mention on this in the results or discussion section.
The variables used in the multivariate analysis were all drawn from the 2008 data. Table 2 shows that there is relatively little missing information in this data. The missing data the reviewer mentions is from the 1994-1997 study. This data was used only to establish comparability between groups. It was not a predictor in the multivariate analysis presented.

Page-12: I would suggest a small heading after line 46, "Cohort or participant description" since this

paragraph describes the combine sample of 381 rather than the individual groups.
Heading added at reviewer's suggestion.

Page-14: It will be helpful to transform back the MABC to its original scale when reporting in the results but either way its fine statistically, just wondering how noon-statistician will interpret it.
The reviewer raises an interesting point about back-transforming the MABC score. We consider this unwise, because the MABC test results are traditionally used to identify probable disability; scores are not meant, as UNIT scores are, to reflect the range of ability found in a normal population. Back transforming the score might encourage readers to over interpret the score, rather than focusing on the between group difference, the main subject of this paper. Although the reviewer rightly wonders what meaning 44% of a standard deviation will have to a lay audience, we consider this the best way to present the result.

Table 2- some of the questions i.e. contemporary SES, presumably the subjects can choose more than one question or are those variable separate, for example, Has Latrine (Yes: 47 (23%) and thus 77% are No). This is not obvious in tables, could the authors make more apparent by putting a foot note or expanding the table to have more rows that show Yes/No.
The footnote the reviewer requests is added to the table.

Well done to authors, the articles read very well and very informative, specially reporting the missing data in the explanatory variables.
Thank you. We appreciate your valuable feedback.

VERSION 2 – REVIEW

REVIEWER	GV Krishnaveni Research Scientist and Wellcome Trust Intermediate Fellow Epidemiology Research Unit CSI Holdsworth Memorial Hospital Mysore, India No competing interests
REVIEW RETURNED	02-Apr-2013

- The author have completed the checklist but made no further comments.