Author's response to reviews

Title: Assessment of data quality in a multi-centre cross-sectional study of participation and quality of life of children with cerebral palsy

Authors:

Heather Dickinson (heather.dickinson@ncl.ac.uk)
Kathryn Parkinson (kathryn.parkinson@ncl.ac.uk)
Vicki McManus (vmcmanus@enableireland.ie)
Catherine Arnaud (carnaud@cict.fr)
Eva Beckung (eva.beckung@vgregion.se)
Jerome Fauconnier (jerome.fauconnier@imag.fr)
Susan I Michelsen (sim@si-folkesundhed.dk)
Jackie Parkes (j.parkes@qub.ac.uk)
Giorgio Schirripa (deceased@deceased.it)
Ute Thyen (thyen@paedia.ukl.mu-luebeck.de)
Allan Colver (allan.colver@ncl.ac.uk)

Version: 2 Date: 29 September 2006

Author's response to reviews: see over
Dear Dr Pemberton

We thank the referees for their careful reading of our paper and for their helpful comments, to which we respond below:

**Reviewer:** David L Streiner  
**Reviewer's report:**  
**General**
A nice illustration of factors affecting response rates

---

**Major Compulsory Revisions**

1. A large number of analyses were run, leading to the problem of some results possibly being significant just by chance. Some correction for multiplicity should be incorporated (e.g., Bonferroni, Holm, Hochberg).

Our analyses are intended to be largely descriptive rather than hypothesis testing. Only 11 p-values are explicitly presented (see 8 p-values presented in Table 3; 3 further p-values presented in the text). We therefore believe that the best way to deal with the potential problems of multiple analyses is:

- To discourage over-interpretation of the findings reported in table 3, we have removed the word "significant" when describing variation between regions (see Results section). We have also added a sentence to the Discussion (see section Heterogeneity between regions on pages 19-20) to discourage over-interpretation of differences between regions, not only because of the problem of multiple analyses, but also because of the correlations discussed in point 2 below.

- We have changed to using a more extreme significance level than the usual p<0.05 criterion: we now use a 1% significance level for all analyses - see changes to Results section of Abstract, 2nd last sentence of Statistical Methods on page 9, Non-response of families selected from population-based registers at top of page 12.

2. Many of the analyses involve variables that are highly correlated with variables in other analyses. For example, country is correlated with socioeconomic status, educational level, and so forth. It is misleading to do separate analyses on these without recognizing that country, for instance, is confounded by SES and education. In order to draw conclusions about country, those other variables must be entered as covariates.

We appreciate that it is not possible to draw conclusions about differences between countries, without allowing for other correlated variables. However, it is very difficult to allow for correlated variables using formal statistical methods in this situation as only 7 countries (9 regions) were represented and a multivariate regression model including several covariates with only 7 (or 9) degrees of freedom could be very misleading. We have therefore been careful not to draw conclusions about differences between countries, and, indeed, it was not the purpose of the SPARCLE study to do so. We have restricted ourselves to highlighting the substantial
heterogeneity between countries, and the consequent recommendation that "all analyses should allow for region as either a fixed or random effect" (see Conclusions).

- We have added a sentence to the section Heterogeneity between regions on page 19-20 to clarify these issues.

Similarly, when SES is examined, country and education should be covaried. Otherwise, there is the impression that each of these variables are independent predictors of compliance. We were unable to consider either SES or education as predictors of non-response, as these variables were not available for non-responders (see section Predictors of family non-response on page 8).

3. The problem with stepwise regression is that variables may not enter because of two reasons: (a) they are not important, or (b) they are important, but are correlated with variables already in the equation. Without examining the full-rank solution, and the correlations amongst all of the predictors, it is impossible to determine which is the case. I would suggest not using stepwise regression.

We apologise: it was incorrect to state that "Forward stepwise regression, with a p-value <0.05 to enter or remove variables, was used to assess whether variables were independently significant." (see section Characteristics of self-completing children). All odds ratios reported in Table 4 are univariate odds ratios.

- Section Characteristics of self completing children has been re-written
- This has now been clarified in the legend to Table 4.

An additional correction has been made to Table 4: as children with learning difficulties were less likely to be able to complete questionnaires themselves, it is likely that IQ was closely related to self-completion. The main point of interest is (i) to confirm that IQ is related to self-completion and, (ii) to assess whether other aspects of impairment are related to self-completion, after allowing for IQ.

- This has now been done - see changes to sections on Characteristics of self-completing children (page 9), Completion by children (page 14), Children’s self-report (page 209) and table 4.

--------------------------------------------------------------------------------------------------

Minor Essential Revisions

Dismissional Revisions

1. The Background says that, because children with CP have a range of impairments, they are "representative of the wider population of disabled children." I would disagree with this; they are not representative of, for example, children with epilepsy, or other more specific disorders. I would suggest dropping this sentence.

- We have now reworded this sentence
2. Because the children in North West Germany were the only ones not part of a register, I would suggest dropping them from the analyses. The numbers are small (n = 75), but the methodological differences between the way they were found and the rest of the sample are large.
A major aim of this assessment of data quality was to assess whether the inclusion of the children in North West Germany, who were not identified from a population-based registry, had the potential to introduce bias. We therefore feel it is important to retain this analysis.

- We have now clarified this by adding a additional sentence to the final paragraph of the Background (see page 4).

Reviewer: Jeanne Landgraf

-----------------------------------------------------------------------------------

Major Compulsory Revisions
-----------------------------------------------------------------------------------

Minor Essential Revisions
-----------------------------------------------------------------------------------

This is a well written article with sound methods and analyses. I have a few minor comments. The article would be greatly strengthened and have wider appeal if a more compelling position were made with regard to the necessity of the work. It seems naturally intuitive that one would perform these analyses as a matter of course when performing any study for which data is collected across multiple countries - or even samples within countries for that matter. Is this not the case? We agree that, ideally, assessment of data quality, non-response and potential for bias should be performed for any epidemiological study. However, as discussed on page 20 (see section Comparison with other studies), we found few reports of such studies.

Far more interesting would be the degree to which sampling bias may empirically affect ratings of QOL or other outcomes measures. The main objective of the SPARCLE study is to estimate associations between QOL (and other outcome measures) and potential explanatory variables. It was not designed to estimate population averages. The degree to which non-response bias may empirically affect such associations will be addressed in papers that report specific outcomes in detail.

Are your findings unique to registries in general? We were unable to identify any other epidemiological studies which identified children from population-based registries and administered face-to-face interviews. Hence we are unable to comment on whether our findings are typical of such registry-based studies.

- We have added a sentence to the section Family non-response on page 16 to clarify this point.
Wouldn’t some of these issues you raise be mitigated by rigorous sampling methods commonly used for large scale epidemiologic studies (eg stratifying by block, neighborhood, age, gender, ethnicity or race etc) for which a registry may not be available.
Stratified sampling is often important to achieve greater precision in estimates of population means; however, the main objective of the SPARCLE study is to estimate associations between factors, rather than means. Stratified sampling is appropriate to prevent selection bias, but it would not have prevented non-response bias, although it would have allowed us to adjust for it more readily. However, in the SPARCLE study, it was not possible to stratify by socio-demographic characteristics, firstly as these were not recorded on the registers and, secondly, because this would have substantially reduced the number of children eligible for the study and hence reduced statistical power.

- This issue is now discussed in an additional paragraph in the section on Family non-response on page 16.

Also, it reads as if the focus is really on the KIDSCREEN as your primary measure - or so it seems - that’s not an issue - its just that this should be clarified in the title - - I am not sure why you mention the other measures and include them in Table 5 when there is no narrative really about them.

Our intention in this paper was to report data quality for all measures - both outcomes and putative explanatory variables - used in the study, as future papers will consider all these measures in structural equation modelling of determinants of participation and QoL. We thank the referee for pointing out that the title is inconsistent with this.

- We have therefore modified the title.
- We have modified the first two sentences of Background on page 4

Page 4 - Background - I am a firm believer that children can self report and your paper is about a self reported measure (KIDSCREEN). (Actually perhaps a bit more description is needed as not all readers will be familiar with this measure).

- As suggested, we have added more detail to the section Interview of children and parents on page 6.

But the 2nd sentence as worded (suggesting that children with CP often have impairments of learning) lays the framework for a justification and argument against this tenent of thought.
We agree that children should self-report if possible. However, we did not want to exclude children with severe learning difficulties from the overall study of the relationship of environment to quality of life. For such children, quality of life necessarily has to be a proxy measure reported by parents.
We realised we had omitted an important analysis. We have made an addition to Table 4: as children with learning difficulties were less likely to be able to complete questionnaires
themselves, it is likely that IQ was closely related to self-completion. The main point of interest is (i) to confirm that IQ is related to self-completion and, (ii) to assess whether other aspects of impairment were related to self-completion, after allowing for IQ.

- **This has now been done - see changes to sections on Characteristics of self-completing children (page 9), Completion by children (page 14), Children's self-report (page 20) and table 4.**

You also state that there are registries available for CP which reduce the risk of selection bias. Yet your study suggests otherwise. Please clarify.

Our study showed substantial non-response associated with specific socio-demographic characteristics; non-response therefore has the potential to introduce bias. However, non-response bias is different from selection bias.

- **This issue is now discussed in an additional paragraph in the section on Family non-response on page 16.**

Finally, SPARCLE is not just about CP but children with any disability, correct? If so then a much a stronger rationale is needed for why CP is the focus of this paper. SPARCLE is restricted to children with cerebral palsy.

- **The first sentences of both the Abstract (page 2) and the main text (page 4) have been changed to clarify this - see page 4.**

There are a lot of conditions where impairment is evident. Shouldn't the same assessments of data quality be made for any study - irrespective of the child's condition?

We agree the same assessments of data quality be made for any study, irrespective of the child's condition. However, as reported as discussed on page 20 (see section Comparison with other studies), we found few reports of such assessments.

Page 6 - why additional questions on pain? What's the rationale?

KIDSCREEN does not include any questions about the child's experience of pain, which is a potential determinant of QoL.

The questions we used are taken from the Child Health Questionnaire.

- We have now clarified this in the section Interview of parents and children (page 6).

Page 6 - what do you mean by "other psychometric measures" in your statement that these were regarded as potential effect modifiers? Further clarification and explication are needed.
By "other psychometric measures", we mean the child's frequency of participation in discretionary activities and behaviour as recorded using the Frequency of Participation and Strengths and Difficulties questionnaires respectively, their experience of pain, and parental stress as recorded using by the Parental Stress Index.

- We have now clarified this in the last sentence of the section Interview of parents and children (top of page 7).

Finally, were you surprised by any of the findings? What are the practical implications for researchers?

The practical implications for researchers are outlined in the last three paragraphs of the Conclusions. We found it surprising that self-completion was related to impairment after allowing for IQ.

- This has now been highlighted in the section on Children's self-report on page 20.