Author's response to reviews

Title: Development of spasticity with age in a total population of children with cerebral palsy

Authors:

Gunnar Hagglund (gunnar.hagglund@med.lu.se)
Philippe Wagner (PW@nko.se)

Version: 3 Date: 13 August 2008

Author's response to reviews: see over
Dear Editor,

We appreciate the careful review of our paper. We have answered each referee separately. Some questions and answers are repeated.

**Answer Referee 1 (Freeman Miller)**

We appreciate the positive response. The typographical error has been corrected.

**Answer Referee 2 (Reidun Jahnsen)**

Comment para 2. We choose levels 1+ and 2 as break-off points for our dichotomization. By having level 1 and 1+ in the same group the discussion regarding the ordinal relationship between level 1 and 1+ (as mentioned in the Discussion para 3) is no problem. As it is sometimes difficult to separate limb stiffness resulting from muscle shortening from stiffness caused by spasticity (as mentioned in the Discussion para 4) we choose not to separate the higher levels in the scale. We have added this into para 4 in the Discussion.

Comment para 4. The placement of some paragraphs could be a matter of discussion. We regard paragraph four and seven in the Methods as a description of the material and an explanation of the method. The first part of paragraph five in the Discussion is a repetition of information in paragraph four in the Methods. We have changed reference to the first sentence in the Background. The misspelling of ITB has been corrected.

Comment para 5. We have added reference to the statement.

Comment para 6. The number in the abstract has been corrected.

**Answer Referee 3 (Kerr Graham)**

Comment para 2. Some studies have shown low reliability with MAS, discussed in paragraph 4 in the Discussion. The low reliability will, in absence of systematic errors, only increase the observed variation in the population spasticity measurements, as well as change their distribution. The consequence of the increased variation to the calculated estimates is simply an increase in confidence interval width. Furthermore, because we have used robust confidence intervals in the analysis, our confidence intervals are rather robust with respect to potentially erroneous specification of measurement distribution. Therefore, in absence of a systematic error, the low reliability should only affect our ability to correctly detect an actual change. A significant finding is likely to be accurate.

The validity of MAS has also been questioned. We have added a description of some recent studies in paragraph 5 in the Discussion. The main concern regarding the validity is that MAS sometimes does not distinguish between resistance caused by spasticity and resistance caused by muscle shortening. As mentioned in paragraph 4 in the Discussion, the decreasing range of ankle motion with age in the total population would result in an underestimation of the decreasing spasticity with age found in this study.

Comment para 3. There are 13 child habilitation centres in the area. The examinations were often done by two physiotherapists examining the child together. In total, about 80 physiotherapists have been involved. This information has been added to the Methods. The
physiotherapists have been educated, by a coordinator in the CPUP-program and by the CPUP-manual.

Comment para 6 The discussion regarding validity and reliability of MAS has been modified according to suggestions from the reviewers.

Comment para 9 The typographical error has been corrected.

Answer question 1 The main reason why we have chosen to dichotomize the Ashworth scale was to make possible the modelling process in which we determine the number of age intervals and their corresponding endpoints. The modelling process in the case with a dichotomy demanded comparison of about 125 model AIC values per diagnosis group. Without dichotomization, this number would increase to roughly $4 \times 125 = 500$, since we would have been forced to use multinomial regression and separate modelling for each category. Proportional odds model was not applicable to our data, because the proportional odds assumption was not satisfied. Furthermore, if not the development of all categories could be described in a piecewise linear fashion, the number of models to evaluate would increase further and interpretation of results would prove difficult. In the included figures it may appear that the piecewise linear assumption would prove valid. However, models are constructed with respect to the logarithm of the odds ratio, which can behave differently than the probability shown in the figures.

Additionally, since the analysis is stratified with respect to diagnosis, we had to find a model with very general assumptions that could be applicable to all groups, in order to get comparable results.

So, the reason why we felt that it was necessary to dichotomize the scale, was that we wanted to estimate the development of spasticity with age with a method that would supply equivalent estimates and confidence intervals for all diagnosis groups. Additionally, we wanted the method to allow for the possibility of a change of magnitude and direction of this development. This was not likely to have been possible with any other method.

Answer question 2 In the CPUP-program both spasticity and passive range of motion is measured. In the present study, we have not analysed the spasticity in relation to ROM.

Answer question 3 Se the comments on para 2 and 3 above.

Answer question 4 Yes, children operated with gastrosoleus lengthening were included, we have clarified that in paragraph four in Methods

Answer Referee 4 (Jan Willem Gorter)

Comments para 3 We have explained the data-analysis in more detail when it comes to the method of differentiating between longitudinal and cross-sectional effects (Methods – para 8)

A similar method to that used in the analysis in order to differentiate between longitudinal and cross-sectional effects, can be found in ”Applied Longitudinal Analysis” by Fitzmaurice et al. (John Wiley & Sons, 2004), page 418 – 421. However, instead of using a mixed model to
model individual development of spasticity with age, we have used a marginal model to model the development of the population as a whole.

The method basically consists of adding to the set of model covariates that estimate development with age, a covariate that accounts for cross-sectional birth cohort effects. This leaves the estimates corresponding to development with age unaffected by cross-sectional effects, such as that of the differences between birth cohorts, and can therefore by considered as estimates of longitudinal age effects.

Concerning the point of the analysis appearing to be data driven, we tried to avoid these type of findings by basing our analysis on an objective criterion. The basic research question is of course whether there is any effect of age on the development of spasticity. In order to answer this question we have to estimate the age effect. If we were to only include age as a covariate in our analysis, no consideration would be taken to the possibility of different development between different ages. Since we felt that this was a real possibility, the analysis had to take this into account. We did so by including linear splines in the logistic regression analysis, which intuitively can be viewed as estimating separate age effects for different age intervals. In order to avoid basing the choice of interval endpoints on data-specific properties, we applied the AIC criterion as an objective measure of the best interval points to use.

In absence of any age effects, the regression estimates for each interval would not be significantly different from zero, especially considering the waste body of data on which the analysis is based.

However, here one could of course argue that a large number of intervals of this kind would allow for false detection of changes in estimates specific to the given dataset. However, this is what the AIC is used for. In a sense, it controls the number of intervals in relation to the available degrees of freedom in the analysis, so that the number of intervals used is meaningful when describing the development with age.

Comments para 4  Children with unilateral spastic CP do not have spasticity in the contralateral leg. Including the contralateral leg would result in a parallel shift of the risk age trend, but the change in increase and decrease with age would not be affected.

The main reason why we have chosen to dichotomize the Ashworth scale was to make possible the modelling process in which we determine the number of age intervals and their corresponding endpoints. The modelling process in the case with a dichotomy demanded comparison of about 125 model AIC values per diagnosis group. Without dichotomization, this number would increase to roughly 4*125 = 500, since we would have been forced to use multinominal regression and separate modelling for each category. Proportional odds model was not applicable to our data, because the proportional odds assumption was not satisfied. Furthermore, if not the development of all categories could be described in a piecewise linear fashion, the number of models to evaluate would increase further and interpretation of results would prove difficult. In the included figures it may appear that the piecewise linear assumption would prove valid. However, models are constructed with respect to the logarithm of the odds ratio, which can behave differently than the probability shown in the figures.
Additionally, since the analysis is stratified with respect to diagnosis, we had to find a model with very general assumptions that could be applicable to all groups, in order to get comparable results.

So, the reason why we felt that it was necessary to dichotomize the scale, was that we wanted to estimate the development of spasticity with age with a method that would supply equivalent estimates and confidence intervals for all diagnosis groups. Additionally, we wanted the method to allow for the possibility of a change of magnitude and direction of this development. This was not likely to be possible with any other method.

**Answer Major Comments para 1** We have included the study by Scholtes et al, and some other recent studies on MAS reliability.

**Answer Major Comments para 2** All measurements were done with the child lying prone with the hip and knee extended. This clarification has been added to the Methods.

The low reliability will, in absence of systematic errors, only increase the observed variation in the population spasticity measurements, as well as change their distribution. The consequence of the increased variation to the calculated estimates is simply an increase in confidence interval width. Furthermore, because we have used robust confidence intervals in the analysis, our confidence intervals are rather robust with respect to potentially erroneous specification of measurement distribution. Therefore, in absence of a systematic error, the low reliability should only affect our ability to correctly detect an actual change. A significant finding is likely to be accurate.

The discussion regarding validity and reliability of MAS has been modified according to suggestions from the reviewers.

**Answer Major Comments para 3** Se above (Comments paragraph 3 and 4)

**Answer Referee 5 (Ilona Autti-Rämö)**

**Answer General comment** Some studies have shown MAS to have low reliability, this is discussed in paragraph 4 in the Discussion. The low reliability will, in absence of systematic errors, only increase the observed variation in the population spasticity measurements, as well as change their distribution. The consequence of the increased variation to the calculated estimates is simply an increase in confidence interval width. Furthermore, because we have used robust confidence intervals in the analysis, our confidence intervals are rather robust with respect to potentially erroneous specification of measurement distribution. Therefore, in absence of a systematic error, the low reliability should only affect our ability to correctly detect an actual change. A significant finding is likely to be accurate.

The validity of MAS has also been questioned. We have added a description of some recent studies in paragraph 5 in the Discussion. The main concern regarding the validity is that MAS sometimes does not distinguish between resistance caused by spasticity and resistance caused by muscle shortening. As mentioned in paragraph 4 in the Discussion, the decreasing range of ankle motion with age in the total population would result in an underestimation of the decreasing spasticity with age found in this study.
We decided to study a total population of children with CP. Using a total population we minimize the risk of selection bias, which would be a problem if using a homogenous subgroup. In analysing the data, consideration was given to the difference between the actual longitudinal age effects of interest, and cross-sectional birth cohort effects.

**Answer question 1** There are 13 child habilitation centres in the area. The examinations were often done by two physiotherapists examining the child together. In total, about 80 physiotherapists have been involved. This information has been added to the Methods. The physiotherapists have been educated, by a coordinator in the CPUP-program and by the CPUP-manual.

**Answer question 2** New physiotherapists were trained by a coordinator in the CPUP-program, and by doing the measurement together with a more experienced physiotherapist. We have not done any test of reliability (see answer General comment above). All measurements were done by the child’s local physiotherapist, well known to the child and family.

**Answer question 3** 89 children were followed from childhood to early adolescence in the study, i.e. from younger than 4 years of age at inclusion up until the age of 11 or further. However, in this type of study there is an underlying assumption of a common development of spasticity with age for all children with the same CP-type. This assumption has to still be present if we were to include only these 89 children in the study, in order to be able to generalize our possible findings. The method of analysis applied in the present study relies on the same assumption. Instead of excluding data to study only the 89 children with full follow-up, it additionally uses children born late to estimate early development and those born early, but were included late, to estimate later development, and so on. This is possible because the method separates change with age from cohort effects, such as different birth-cohorts having different underlying risks from being exposed to different methods of treatment. We can therefore be satisfied that we are indeed estimating development with age and not something manufactured by late inclusion and/or early dropout of children from different birth cohorts.

**Answer question 4** According to SCPE dyskinetic CP is defined as: “involuntary, uncontrolled, recurring, occasionally stereotyped movements. Primitive reflexes pattern predominate, muscle tone varies”. In the thesis by Kate Himmelmann “Cerebral palsy in western Sweden”, Gothenburg 2006, 69% of children with the dystonic type of CP had signs of spasticity. Accordingly children with Dyskinetic CP may have, and often have, spasticity.

**Answer question 5** The children are measured twice a year up to 8 years of age, then once a year.

**Answer question 6** Table 3 shows the number of measurements, not the number of children. The number of children in the total material was almost equal in each age-group 3-9 years of age.

**Answer question 7** We do not treat children with CP with electrical stimulation. We use night splints and serial casting for treatment of muscle shortening, but not for treatment of spasticity. Serial casting and night splints are mainly used in the younger ages. If these treatment methods for muscle shortening were to have any influence on the child’s level
of spasticity, it would consequently strengthen the result, in the form of a more pronounced decreasing spasticity in the older age groups.

Answer question 8 It would be an interesting analysis, but GMFM is not included in the CPUP-program