Author's response to reviews

Title: Validation of the Diagnosis of Autism in General Practitioner Records

Authors:

Dr Eric Fombonne (eric.fombonne@mcgill.ca)
Lisa Heavey (nunsnunsreverse@aol.com)
Dr Liam Smeeth (liam.smeeth@lshtm.ac.uk)
Laura C Rodrigues (laura.rodrigues@lshtm.ac.uk)
Ms Claire Cook (claire.cook@lshtm.ac.uk)
Peter G Smith (peter.smith@lshtm.ac.uk)
Linyan Meng (linyan.meng@muhc.mcgill.ca)
Prof Andrew J Hall (andy.hall@lshtm.ac.uk)

Version: 2 Date: 5 Feb 2004

PDF covering letter
Dear Editor,

You will find enclosed the revision of our manuscript and a list of changes that were made in response to the reviewers’ comments. The reviewers’ comments appear in italics and our response to each of them in normal characters.

Reviewer 1

· Abstract – The mean age of diagnosis is not the primary finding and need not be reported in the Abstract and certainly should not come first in the Results section.

We have moved this sentence later in the Results section of the Abstract in response to the reviewers’ comment. We consider it is still useful information to keep in the Abstract to describe our sample.

· Abstract – The kappa for agreement of a PDD diagnosis is on the boundary of moderate to good to my mind. It might be helpful in the text for the authors to present the number of cases on which there was not agreement.

Did one rater rate more cases as PDD than the other or did both rate approximately the same proportion of cases but disagree more often? I am assuming this figure was for the global clinical judgment of PDD presence/absence but such an agreement could also be calculated for the algorithm diagnosis.

Armitage and Berry (Statistical Methods in Medical Research, 3rd ed, 1994, Oxford) describe a kappa of 0.75 or more as indicating “excellent agreement” so we think it is reasonable to describe a value of 0.73 as indicating “good” agreement. We have changed ‘high’ to ‘good’ in the Abstract, and this is consistent with what was originally said in the text under Interrater reliability. In that section, we have indicated that disagreement occurred in only 2 cases which is excellent.

· Page 3, para 1 – The second sentence appears to argue for a causal association from increased research activity in epidemiology leading to refinement in the definition of autism. This is not the case. The third sentence says that the definition ‘now involves’ a specific combination of symptoms – Has it not always done so?

We agree with the critique and have rephrased this sentence correspondingly.

Page 4, para 3 – Spell out VAMP.

It is now spelt in the text.

· Page 5, para 1 – Spell out READ.

Read is the name of the person who devised the READ coding system. We therefore left it as such in the text.
Can the authors clarify what information was and was not available on the N=38 cases that did not have complete records (and were thus excluded).

We did not obtain the records on 366-318=48 (not 38) subjects. No information was available on these 48 children as we did not obtain the medical reports.

Clarify that it is onset or emergence of words and phrases that is intended.

It is not clear what distinction the reviewer has in mind. We coded the age of first words and first phrases as it is usually referred to and recorded in medical records. It was not possible to make more fine grained differentiations in language development for obvious reasons.

One alternative more conservative approach would be to use both the adjusted DSM algorithm and the stricter full algorithm. The data on agreement between clinical judgment of the 2 raters and algorithm diagnosis is not presented but it could be for this 3-way analysis: strict algorithm vs. looser algorithm vs. clinical judgment.

The method of record review used to confirm “caseness” in our study makes it difficult to apply the stricter full algorithm. We did not provide kappa statistics between the adjusted DSM algorithm and the clinical judgement since they were not independent. The kappa is provided for the agreement between the 2 clinical ratings on a subset of 50 patients as they were independently derived.

The authors should note that this analysis is not independent of autistic disorder vs. PDD/AS previously conducted.

What is meant here is not entirely clear. We explain in the text that we explored the correlates of regression only in the Autism group as there were only 3 cases in the PDD/AD group who had regression (see Table 2). Thus, we restricted the next analysis that aimed at identifying the characteristics of regressive and non-regressive cases to the Autism group since it was both a homogeneous diagnostic group (having kept the PDD/AD in the analysis would confounded the results) and a group with sufficient numbers of regressive cases.

I am not sure if the authors have made the point about clinical presentation becoming less severe over time clearly. My understanding is that it is likely that over time more children with less severe presentations are receiving autism or PDD diagnosis in the GPRD but we do not know if the presentation for cases (more severe) that would previously have been included has changed. I think that Table 10 might work better in terms of readability and subgroup sizes to group the cases in 10-year intervals (at least for the 2 oldest cohorts).

It seems that we probably did not explain well why we looked at trends over time in indicators such as % of male gender, % of mental retardation, and % of epilepsy. All these indicators are robustly associated with increased severity of autism in the epidemiological and clinical literature on autism. Thus, if there were convergent changes in these indicators over time in the GPRD database, that would suggest decreasing (or increasing) levels of severity. This is exactly what we found. We have now slightly reformulated this paragraph to make it more explicit. We however kept the 5-year grouping as it allows us to have equal intervals over the 25 years of the study and it provides a more accurate assessment of the trend. We would have lost this accuracy.
and it might have been confusing for the reader if we had grouped the data in 10 year intervals for the 2 oldest cohorts and in 5 year intervals for the most recent birth cohorts (as was suggested by the reviewer).

· Page 11 – Can the authors comment on the 2 variables where there was less reliability: regression and intellectual functioning.

The assessment of regression was less reliable due to the difficulty to operationalize regression in the course of language development and to differentiate proper regression (with loss of skills) from developmental stagnation. Secondly, the kappa statistic is vulnerable to low base rates and regression was relatively infrequent. We have added a sentence to explain this in the text.

For intellectual functioning, the figure of .57 corresponds to the kappa calculated for the agreement on IQ coded on six levels (normal, low average, mild MR, moderate MR, severe MR, profound MR). Considering the small sample, we have grouped IQ in 3 bands (normal range, mild MR, moderate to profound MR). Based on this, the kappa is 0.72. We have now included this figure in the manuscript alongside an explanation of the meaning of the 3 IQ levels.

· Page 12, para 2 – Report only % and not (repeat) numbers in the text.

We have not changed this. The numbers in brackets refer to studies in the reference list, and not to sample sizes. We felt it useful to provide the actual % of regression reported in each study quoted in this paragraph as this is very useful information to have for the readership.

· Page 13, para, line 5 – I think the authors mean children with atypical autism (note this term has not been previously used in this MS) did not get a diagnosis of PDD in this study analysis but they might mean in the GP records – Can they clarify? I do not follow the argument made in this paragraph about a decrease in severity of autism as indexed by changes over time in gender, epilepsy, IQ and language necessarily supporting their first explanation. First, rather different effects might be operating. For example, epilepsy might increase with age and thus be higher in older cohorts. Alternatively, phrase speech might decrease due to the younger age at which children are seen. This paragraph was one of the few (in fact the only) one where I could not follow the logic of the argument.

We have clarified why we looked at changes in clinical characteristics of autism over time (see above, page 10). We had already mentioned in the text that the interpretation of trends was made difficult due to age differences particularly for language level. We agree with the reviewer that this applies as well to rate of epilepsy and we have accordingly rephrased the text. The robust upward trend in % of males and % of normal IQ cannot, however, be confounded by age and illustrates the fact that severity has decreased. As mentioned in the text, several explanations for this trend must be contemplated.

Reviewer 2

This reviewer made no other comments than to express interest for the work and recognize its usefulness in the overall context of our case-control study.
Reviewer 3

1. The question(s) addressed are either not new or otherwise focused.

2. The method for evaluating the validity of the diagnosis of autism in the GPRD is appropriate and reasonably well described. However, work on this subject, i.e., validity of diagnosis of autism in the GPRD, has already been published (but not referred to in the current manuscript) – Black C et al. BMJ 2002;325:419-21. The methods for describing various characteristics of the cases of autism are unsatisfactory and likely misleading since the case group is completely heterogeneous, i.e., consists of males and females ages 0-36+ and different calendar time periods. Since the epidemiology of autism is likely to vary in substance among or within these groups, summary statistics for these categories combined are at best uninterpretable and at worst misleading.

3. The data are reasonably sound but not properly “controlled.”

4. The tables of the data are superficial and provide little if any substantive information.

5. The conclusion that the validity of the diagnosis of autism in the GPRD is high is valid. The remaining “conclusions,” such as they are, are at best not useful.

6. The title and abstract accurately convey a part of what is in the manuscript.

7. The writing is adequate.

Specific Comments

1. The first sentence of the abstract is out of place. One expects to get the results of a “large case-control study,” which, of course, is not what the authors provide.

2. The inclusion of considerable space and attention to “PDD” confuses the issue relative to autism itself which is the main focus of the study.

3. The description of the GPRD is entirely inadequate for most readers. Reference 21 is not the appropriate reference for the preceding sentence. Also, the former MCA is currently called the MHRA (Medicines and Healthcare Products Regulatory Agency).

4. Much of the results provided in the tables are misleading. For example, in Table 1, the “mean age” of language level recording is given as 7.86 years. Only informed readers will be aware that a substantial majority of cases of autism are diagnosed prior to age 6. Age of parental recognition is apparently given in “years” although at first glance one might assume it is months. Age at first diagnosis of autism is given as a mean of 5.45 years. This is a highly misleading statistic to use since the median is far younger. Table 4 is highly misleading. The quality and extent of the available information in the GPRD varies greatly over these calendar time periods.

5. The authors can safely wait until their “case-control” study is published to briefly present their results relative to the validity of the diagnosis of autism.
6. This manuscript is a reflection of the absence of the authors’ experience in the use of the GPRD.

The result is a superficial and inadequate presentation of results provided in the manuscript.

The assessment from this reviewer contrasts markedly with that from the other 2 reviewers. We are concerned that the nature of this review may have been influenced by other factors that we think it is important to bring to your attention.

We wish to disclose to you correspondence that was exchanged recently between the Editor of the BMJ, this reviewer and one of us. (copy of correspondence enclosed). In this correspondence, reviewer 3 wrongly attributes a review which lead to a manuscript of his to be rejected to one of us (Dr Hall) and Dr Jick accused Dr Hall of a lack of professionalism. Following a response from Dr Hall, Dr Jick recognized that Dr Hall had not been the reviewer and apologized, but we think it likely that he reviewed our manuscript while under this mistaken impression. The tone and nature of his review suggests to us that he may not have been entirely objective in his comments. In view of this experience we request that Dr Jick has no further role in the review of our paper.

In reviewing his comments, we consider that his remark on our omission of his previous validation exercise on the GPRD database was fair and we now cite this study in the text (first sentence of the discussion) and in the reference list (now study 25).

Sincerely,

Eric Fombonne et al.