

The prevalence of autism spectrum disorders: impact of diagnostic instrument and non-response bias

M. Posserud · Astri J. Lundervold ·
Stein Atle Lie · Christopher Gillberg

Received: 18 November 2008 / Accepted: 8 June 2009 / Published online: 24 June 2009
© Springer-Verlag 2009

Abstract

Background A large part of the variability in rates of autism spectrum disorders (ASD) across studies is non-aetiologic, and can be explained by differences in diagnostic criteria, case-finding method, and other issues of study design.

Aim To investigate the effects on ASD prevalence of two methodological issues; non-response bias and case ascertainment. We compared the findings of using a semi-structured parent interview versus in-depth clinical assessment, including an ASD specific interview. We further explored whether including information on non-responders affected the ASD prevalence estimate.

Method A total population of 7- to 9-year olds ($N = 9,430$) was screened for ASD with the autism spectrum screening questionnaire (ASSQ) in the Bergen Child Study (BCS). Children scoring above the 98th percentile on parent and/or teacher ASSQ were invited to participate in the second and subsequently in the third phase of the BCS

where they were assessed for ASD using the Development and Well-Being Assessment (DAWBA), and the Diagnostic Interview for Social and Communication disorders (DISCO), respectively.

Results Clinical assessment using DISCO confirmed all DAWBA ASD cases, but also diagnosed additional cases. DISCO-generated minimum prevalence for ASD was 0.21%, whereas estimated prevalence was 0.72%, increasing to 0.87% when adjusting for non-responders. The DAWBA estimate for the same population was 0.44%.

Conclusion Large variances in prevalence rates across studies can be explained by methodological differences. Both information about assessment method and non-response are crucial when interpreting prevalence rates of ASD.

Keywords Autism spectrum disorders · Prevalence · Non-response bias · Diagnostic instrument · Assessment method

M. Posserud (✉)
Centre for Child and Adolescent Mental Health, Unifob Health,
Jonas Lies vei 91, 5009 Bergen, Norway
e-mail: maj-britt.posserud@rbup.uib.no

A. J. Lundervold
Department of Biological and Medical Psychology,
Faculty of Psychology, University of Bergen,
Jonas Lies vei 91, 5009 Bergen, Norway

S. A. Lie
Centre for Child and Adolescent Mental Health, Unifob Health,
P.O. 7800, 5020 Bergen, Norway

C. Gillberg
Child and Adolescent Psychiatry, University of Gothenburg,
Kungsgatan 12, 411 19 Göteborg, Sweden

Introduction

The traditional triad of autistic behaviours includes social impairment, communication problems, and repetitive behaviour/interests. In the past decades, the concept of autism has evolved from Kanner's narrowly defined severe autism to encompass social impairment across all levels of intellectual functioning. The recognition that the social difficulties characteristic of autism spectrum disorders (ASD) can appear in many different variants depending on, e.g. language skills, general level of ability, severity of symptoms, context, and coexisting disorders, has led to a dramatic rise in the amount of children diagnosed with ASD. This worldwide rise in ASD prevalence has fuelled

debates as to the aetiology of this sudden increase, with vaccines as the easy-to-get culprit in media and public debate. With prevalence data being treated as “absolute truths”, vaccines have been put forward as a possible explanation for the “autism epidemic” [33]. Several rigorous studies have, however, failed to show any link between vaccines and ASD [8, 16, 24], and the prevailing view is that change of diagnostic criteria and increased awareness explain most of the apparent rise in ASD prevalence [6, 10, 34]. Several reviews of ASD prevalence have also suggested that differences in methodology and case-finding can account for a large part of the great variability of ASD prevalence across studies in time and location; the more thorough examination of the study population and the wider the criteria applied, the higher the prevalence [2, 32, 36].

There is then ample evidence to suggest that assessment method and diagnostic procedure have severe impact on the prevalence estimate obtained in a study. Two Norwegian studies examining prevalence of ASD arrived at vastly different results with 0.05% (95%CI not given) in 1998 [29] and 0.44% (95%CI 0–0.91%) in 2007 [15], suggesting an almost tenfold increase in ASD prevalence in Norway in less than 10 years. The studies differed greatly both in regard to case-finding and estimation procedures. Also, awareness of high-functioning children with autism has increased considerably between 1998 and 2007 in Norway. Therefore, it is difficult to know if the differences between these studies reflect any true rise in ASD prevalence. The first study only gathered information by asking other centres to refer suspected cases, whereas the latter study used the standardised parent Development and Well-Being Assessment (DAWBA) [12] interview on a large subsample of a population to diagnose ASD. Using standardised instrument across studies increases the possibility to understand prevalence differences in terms of the aetiology of disorders. The DAWBA is currently available in 16 languages (www.dawba.com), which makes it available to an unprecedented number of people across the world. The cost-effective diagnostic procedure and the availability of the DAWBA also simplify epidemiologic studies. An important limitation of the DAWBA is the lack of validity for the diagnoses arrived at. Validity studies of lay-administered semi-structured interviews on psychiatric problems in adult general populations have shown that the agreement compared to clinical assessment is poor [4, 5]. Although we know of no study on the validity of a DAWBA-generated ASD diagnosis, several studies stating the prevalence of ASD generated from the DAWBA have already been published [7, 15, 25]. These studies may well be comparable to one another, but we know little about the validity of the DAWBA ASD diagnoses. As we know that the case-finding method is important for the resulting

prevalence estimate, it is important to examine how a DAWBA ASD diagnosis compares to the gold standard of a full clinical assessment using autism specific measures.

Another important issue that may affect prevalence estimates is the effect of non-response. The few studies examining non-responders in epidemiologic studies of mental health all conclude that non-responders have a higher prevalence of problems than responders, and also seem to have more severe symptomatology [13, 18, 23]. Failure to account for non-responders would thus lead to underestimation of prevalence and perhaps also underestimate the relationship between disorders and risk factors, as the cases examined have less severe disorder [14]. Although the same pattern of non-response is assumed to appear in studies of *child* mental health [28, 31], few studies have managed to compensate for the effects of this bias on prevalence estimates, and to date, no study has examined non-response in relation to ASD as primary outcome. The information available on non-responders is usually only background variables such as gender, socioeconomic status and geographic location. These background variables, although related to prevalence of various mental health problems, cannot be used to estimate prevalence, but only to estimate the direction of the difference. One study concluded that the net result of selection bias is negligible, because of small effect sizes in mean differences between non-responders and the entire group [9]. However, differences in mean score of a test may be of limited value, given that interest is often focused on the extreme scorers, either low or high. Our group recently showed that even though the differences in mean measured by effect sizes were small, being a high-scorer either on symptoms or degree of impairment, predicted non-response [31].

Aims

The overall aims of the present paper were to (1) estimate prevalence of ASD, (2) validate the DAWBA ASD diagnoses, and (3) to explore the impact of non-response on ASD prevalence estimates. In the Bergen Child Study (BCS), we obtained autism-relevant information at three different levels on the same background population, each with varying degrees of detail. The first phase was questionnaire data (autism spectrum screening questionnaire—ASSQ), with 97% coverage of the entire population. Second phase consisted of a parent interview, the DAWBA, which also generates ASD diagnoses, covering 11% of the total population. Third phase consisted of a full in-depth clinical assessment including autism diagnostic interview and observation, resembling the golden standard of ASD diagnostics, but covering only 3% of the population. In the present study, we used the data from the clinical

assessment to estimate ASD prevalence. Using the questionnaire data from the first phase (covering 97% of the population), we ask whether there is a selection bias from the first phase to the third phase, and if/how this would affect prevalence estimate. We further explore the agreement between the DAWBA and the clinical assessment. Finally, we examine how the ASD prevalence is affected by the use of DAWBA vs. clinical assessment as diagnostic method.

Method

The BCS—first phase

The screening for ASD was part of a larger longitudinal study—the BCS—assessing mental health among children in Bergen. The initial wave of the BCS consisted of three phases. We invited a total population of children to participate in the first phase by sending a questionnaire consisting of the strengths and difficulties questionnaire (SDQ) [11], the ASSQ and other questions to parents and teachers of all children aged 7–9 years in Bergen ($N = 9,430$). The *identified ASSQ sample* in this study consisted of the children whose parents gave informed consent to participate and had a matching teacher questionnaire ($N = 6,609$). The identified sample covered 70% of the population. Teacher questionnaires were filled in for all children, but for children without informed consent the forms contained no information that could be traced back to the individual child (such as age, name, school, teacher etc.). The teacher ASSQ sample covered 97% of the population [26]. The *anonymous ASSQ sample* consisted of the teacher ASSQ without matching parental consent ($N = 2,543$).

The Regional Committee of Ethics on Medical Research in Western Norway approved the study.

Screening criteria

The *ASSQ screen positive criteria* were defined as an ASSQ score above the 98th percentile on either parent and/or teacher ASSQ (i.e. scores exceeding 18 and 15 points, respectively). According to these criteria, 212 children (3.2% of those who had both a parent and a teacher ASSQ completed) were defined as screen positive. There were several other screening criteria to detect other disorders; i.e. scoring above the 90th percentile on the SDQ symptom score on parent and/or teacher questionnaire, scoring moderate/high on the SDQ impact scale, and/or scoring above the 98th percentile on other symptom scales (such as a tic scale) in the BCS questionnaire [15]. For the purpose of this study, these other criteria are referred to as the *overall screen positive criteria*.

BCS—second and third phases

The second phase of the BCS comprised the DAWBA parent interview. Parents of all screen-positive children (either on the ASSQ screen and/or overall screen) and a random 15% of screen-negative children were invited to the DAWBA [15].

The third phase consisted of a broad clinical assessment of mental health and cognitive function with Kiddie-SADS [19] (www.kiddiesads.com), WISC-III, and additional tests. The following four groups of children from the second phase were invited to the third phase: (1) all ASSQ screen-positive children, (2) all children with a DAWBA diagnosis (of any kind, not exclusively those with an ASD diagnosis), (3) a random selection of screen-positive children (overall screen), and (4) a random selection of screen-negative children. Children with a chronic disorder (physical) from the first phase were invited directly to the third phase to increase their participation [17]. A total of 297 children from the identified ASSQ sample came to the full assessment in the third phase of the BCS (Fig. 1).

The ASD study sample

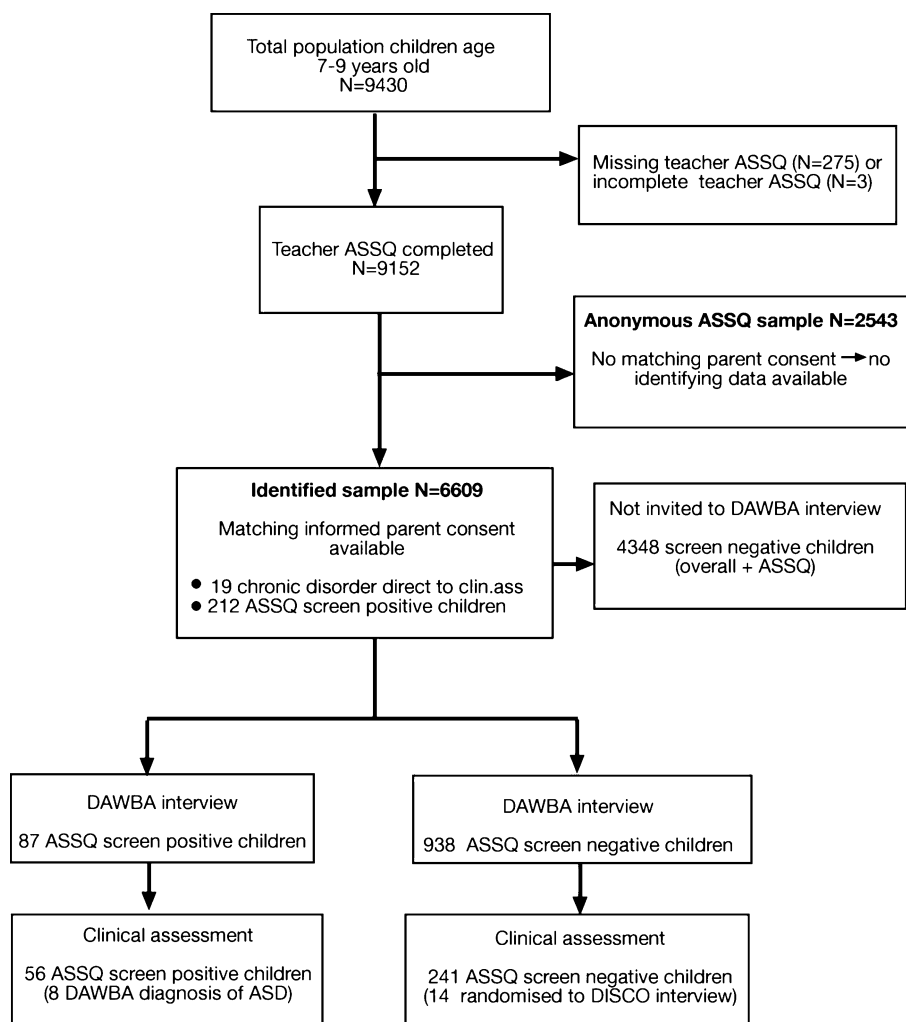
All ASSQ screen-positive children in the identified ASSQ sample were invited to the DAWBA ($N = 212$), and 87 interviews were completed. The ASSQ screen-positive children were all invited to the third phase, regardless of the DAWBA results. In addition, two ASSQ screen-positive children came directly from the first phase as part of the chronic disorder children (see Fig. 1, and Hysing et al. 2007). In total, 56 ASSQ screen-positive children were assessed in the third phase, corresponding to 26.4% of all ASSQ screen-positive children from the identified sample.

ASD assessment tools and diagnostic procedure—second phase

Prevalence estimates for autism, Asperger's disorder and PDD–NOS generated from the DAWBA have been published elsewhere [15]. A total of 1,025 DAWBA interviews were performed, 87 on ASSQ screen-positive children, and 938 on ASSQ screen-negative children. The later versions of the parent DAWBA comprise initial screening questions on early development (seven items), friendships (eight items), and social competence (ten items). If problems in these domains are identified, the DAWBA algorithm opens up a section with up to 36 ASD-relevant questions that, together with four open-ended questions on impact, daily life, function, and contact with health services, form the basis for making a DAWBA diagnosis of ASD.

As the DAWBA does not diagnose the broader autism phenotype (BAP), we calculated kappa agreement between

Fig. 1 Flowchart of ASSQ screen-positive children in the BCS



the DISCO/Kiddie-SADS and the DAWBA for DSM-IV ASD excluding BAP.

Assessment tools and diagnostic procedure—third phase

A certified interviewer (M. Posserud) assessed 48 of the 56 ASSQ screen-positive children in the third phase with the Diagnostic Interview for Social and Communication disorders (DISCO), tenth revision [35]. The remaining eight ASSQ screen-positive children were not interviewed with the DISCO due to the parent/child not wanting to come back for assessment after the Kiddie-SADS interview ($N = 1$) or it was impossible to arrange another interview day ($N = 7$). In all these cases, records were reviewed closely by the first author, and, in addition, either the child was seen by the first author or discussed with the Kiddie-SADS interviewer, who had evaluated both the parent and child during the interview. To ensure finding possible false screen-negative cases, all Kiddie-SADS interviewers were instructed to refer children with suspected social difficulties on the basis

of the Kiddie-SADS results to the first author. In this way, two children, not belonging to the ASSQ screen-positive group, were referred for DISCO assessment. To reduce interviewer bias, a random selection of ASSQ screen-negative children ($N = 14$) were assigned for DISCO assessment, and all DISCO interviews were conducted blindly to the DAWBA results and the ASSQ score. In total, 64 DISCO interviews were completed. The first author also saw and interviewed or interacted with all children whose parents took the DISCO interview, and in 48/64 children she also performed the Kiddie-SADS interview.

The DISCO is an interview for systematic gathering of information that enables the interviewer to make a diagnosis within the autism spectrum [20, 35]. Similar to the Autism Diagnostic Interview-Revised (ADI-R) [22] it involves a diagnostic scoring algorithm to produce diagnoses according to the ICD-10, DSM-IV or other diagnostic criteria, but in contrast to the ADI-R the interviewer is expected to score all the items on the basis of all available information, including observation of the child. Using the DISCO is therefore conceptually similar to a

combined diagnostic procedure involving both ADI-R and the Autism Diagnostic Observation Schedule (ADOS) [21]. In addition, the DISCO contains items on early development and on activities of daily life and thus gives the interviewer some idea of the level of functioning in several different aspects of daily life, not only social functioning and communication.

Diagnoses

For the purpose of this study the concept of ASD comprised Autistic Disorder/Childhood Autism, Asperger's Disorder/Syndrome and Pervasive Developmental Disorder—Not Otherwise Specified/Atypical Autism (DSM-IV, ICD-10). Children with autistic traits but with good functioning or sub-threshold problems were classified as BAP.

Mental retardation (MR) was defined as a WISC-III full scale IQ score (FSIQ) <70, and level of MR (mild, moderate, severe) was defined according to the DSM-IV-TR [1].

All the information from observation, testing, Kiddie-SADS and DISCO was used to make the final clinical ASD diagnosis. All the cases and their protocols were discussed individually in detail with the last author (C. Gillberg) before final diagnosis was made.

Analyses

Minimum prevalence was calculated as the number of ASD cases divided by the number of children in the population they were drawn from (the *identified sample*, $N = 6,609$). *Estimated prevalence* was calculated using the positive predictive value (PPV) of the ASSQ in the third phase, i.e. the rate of ASD among the ASSQ screen-positive children [27]. We multiplied the PPV with the number of ASSQ screen-positive children in the first phase, for parents and teachers independently. Having only teacher ASSQ in the anonymous sample, we estimated the prevalence here by adding the proportion of the ASD cases that were identified by parents only among the identified children. Confidence intervals were calculated from the standard error among the children examined in the third phase.

The rate of ASD among ASSQ screen-negative children (i.e. the false negative rate) could not be used for the prevalence estimate. The BCS study was designed to maximise sensitivity for all different mental health disorders, with broad overall screening criteria and invitation of all screen-positive children (both ASSQ screen and overall screen). Therefore, most ASSQ screen-negative children in the third phase were screen-positive on the overall screen and, thus, not representative for the ASSQ screen-negative children in the first phase. This sampling bias does not apply to the ASSQ screen-positive children, as they were all invited to participate.

Non-response

To evaluate the effect of non-response as a bias on prevalence estimate, we first examined which variables were independently correlated to an ASD diagnosis. We performed logistic regression with ASD as the dependent variable, and SDQ total score, impact score, and ASSQ score from parents and teachers as independent variables. SDQ total score and impact score were selected in addition to ASSQ score because they formed part of the overall screening criteria. The children in the third phase were then compared with children in the first phase using two-tailed t tests of mean scores. The children in the third phase were considered representative of the children in the first phase if there were no differences in mean scores on variables predictive of an ASD.

Results

The DAWBA interview and ASD

Ten of the 87 ASSQ screen-positive children interviewed in the second phase were diagnosed as having an ASD vs. none among the ASSQ screen-negative children. Eight of the DAWBA ASD diagnosed children came to the third phase (two children were non-responders in the third phase), and all eight were independently diagnosed with ASD using the DISCO interview (Table 1). In total, 273 children were seen *both* in the second and third phases. The

Table 1 DAWBA and DISCO diagnoses

Gender	DISCO	DAWBA	IQ
Boy	ASD	ASD	Normal
Boy	ASD	ASD	Normal
Boy	ASD	ASD	Normal
Boy	ASD	ASD	Normal
Girl	ASD	ASD	Mild MR
Boy	ASD	ASD	Mild MR
Boy	ASD	ASD	Moderate MR
Boy	ASD	ASD	Severe MR
Boy	ASD	Negative	Normal
Boy	ASD	Negative	Normal
Girl	ASD	Negative	Normal
Boy	ASD	ODD	Normal
Boy	ASD	ADHDco + soc. phobia	Mild MR
Boy	ASD	Non-responder	Moderate MR
Boy	Non-responder	ASD	Non-responder
Boy	Non-responder	ASD	Non-responder

ODD oppositional defiant disorder, ADHDco attention deficit/hyperactivity disorder combined type, soc. phobia social phobia

kappa agreement for any ASD between the parent DAWBA and the DISCO/Kiddie-SADS assessment (excluding BAP) was 0.75.

The DISCO interview and ASD

Fourteen children were found to have a DISCO-diagnosed ASD; six children were diagnosed with autism (one girl), six children with Asperger's disorder (one girl), and two with atypical autism/PDD (Table 1). In addition, nine boys were classified as BAP. These children had the same typical difficulties, but with less severe symptoms, and functioning well enough to disqualify them for a clinical diagnosis of current ASD.

Five children were diagnosed with an ASD with the DISCO, but did not receive an ASD with the DAWBA, (one child receiving a DISCO diagnose did not attend the DAWBA). These children were characterised by normal intelligence (all except one child with MR). In two cases, observation was crucial for diagnosis, as parents could not give an appropriate account of their child's difficulties. In one child the present difficulties were less obvious, although the developmental history was clearly deviant. One child was very passive, without flagrant ASD symptomatology, and the child with MR was diagnosed as having ADHD combined type and social phobia according to the DAWBA.

Non-response

Both parent ASSQ and teacher ASSQ total scores were independent predictors of an ASD in the logistic regression (variance explained = 0.47, parent ASSQ $P < 0.001$ and teacher ASSQ $P = 0.002$), whereas neither parent nor teacher impact score or SDQ total symptom score predicted ASD.

When examining the mean ASSQ score among the ASSQ screen-positive children, there were no statistically significant differences between the third phase and the first phase (identified children). Both parent and teacher ASSQ screen-positive children in the third phase had similar ASSQ scores as the ASSQ screen-positive children in the first phase (Table 2). The ASSQ screen-negative children in the third phase, however, had statistically significantly

higher ASSQ scores than the ASSQ screen-negative children in the identified sample on both parent and teacher questionnaires.

Prevalence estimates

Minimum prevalence for all ASD was 0.21%. However, since only 297 children were examined, this should be regarded as an absolute minimum figure. *Estimated prevalence* was based on the PPV of the ASSQ among the ASSQ screen-positive children only, as the ASSQ screen-negative children were not representative of the population. Estimated prevalence in the *identified ASSQ sample* was 0.72% (95%CI 0.36–1.07). Estimated prevalence in the *anonymous sample* was 1.28%. Adding the estimated prevalence in the *identified sample* and the *anonymous sample*, the estimated prevalence in the total population was 0.87%.

Discussion

Prevalence estimates of ASD

In the current study, we found that prevalence of ASD varied from 0.21% (minimum prevalence), increasing to 0.72% when extrapolating to all identified children eligible for the study, and further to 0.87% when estimating the extra load that non-responders would have contributed if they had participated in the study, due to their higher rate of teacher ASSQ screen-positive children. This is almost twice as high as the previous study in the same population that arrived at a considerably lower overall prevalence estimate of 0.44% (including non-responders) based on the parent DAWBA. The reasons for this discrepancy are discussed below.

DAWBA versus DISCO ASD diagnosis

The good agreement between the clinical diagnosis in the third phase and the DAWBA diagnosis in the second phase demonstrates that a DAWBA diagnosis of ASD is valid. However, the additional cases found in the third phase also show that the DAWBA misses a substantial part of ASD

Table 2 ASSQ mean score in phase 3 and phase 1 of the BCS

	Parent ASSQ sc+		Teacher ASSQ sc+		ASSQ sc–	
	Parent ASSQ	Teacher ASSQ	Parent ASSQ	Teacher ASSQ	Parent ASSQ	Teacher ASSQ
Phase 1 ($N = 6,609$)	24.4	8.7	13.7	21.1	2.8	1.3
Phase 3 ($N = 297$)	24.5	12.3	12.6	22.1	4.5	3.0
P -value (difference)	0.93	0.13	0.52	0.31	<0.001	<0.001

cases. Thus, one cannot exclude the presence of ASD on the basis of the parent DAWBA interview. The findings of the present study suggest that the DAWBA may have problems recognizing ASD in children within normal range of intellectual function. Four out of five cases missed by the DAWBA had IQ-levels in the normal range, whereas only four out of eight children for whom the DAWBA and the DISCO agreed had an IQ-level in the normal range. Although the DAWBA has a rather extensive section on ASD to go through if there are problems recorded on screening questions, it can hardly replace the extensive assessment that is usually required for a definitive diagnosis of ASD. The DAWBA focuses on present day symptoms, whereas a clinical diagnosis of ASD usually requires a detailed developmental history. The quality of a DAWBA diagnosis also depends on the information provided by the informants on well-defined questions, and correct diagnosis therefore requires informants to understand the questions and to be able to answer them. Parents with ASD traits may have difficulties appreciating qualitative social problems or give adequate narratives of them. Our results indicate that the DAWBA may be adequate in order to diagnose or indicate autism in clear-cut cases with definitive cognitive impairment and in cases where parents and teachers are well aware of the child's difficulties. In children with normal intellectual functioning, uncertain cases, borderline cases, atypical cases and cases where parents fail to acknowledge the difficulties or have problems communicating them, the DAWBA is probably inadequate to diagnose or even indicate the presence of an ASD.

ASD prevalence estimates

In the overall prevalence estimate for ASD of 0.87% we did not take into account that there were two cases of ASD among the ASSQ screen-negative children. The true prevalence is therefore likely to be even higher, but it is difficult from our data to estimate the relative contribution from the false-negative cases to the total ASD prevalence. The two ASD cases among the 241 ASSQ screen-negative children in the third phase would correspond to an ASD prevalence of almost 1% among the ASSQ screen-negative children. However, most of the ASSQ screen-negative children in the third phase were either screen-positive on the overall screen or had a DAWBA diagnosis. Both these groups have a much larger likelihood of containing ASD cases than a randomly selected population of ASSQ screen-negative children. The fact that there were no false-negative children among the 941 children examined with the DAWBA supports the notion that the number of false-negative children should be lower than the data from the third phase would indicate. We could also show that the

ASSQ screen-negative children in the third phase differed significantly on both parent and teacher ASSQ, showing that they were not representative for the general population of ASSQ screen-negative children.

Our prevalence estimate for ASD is considerably higher than previous Norwegian studies showing a prevalence rate for all ASD of 0.05 [29] and 0.44% [15], respectively. The latter study had the same background population as the present study, but was based on case-finding according to the DAWBA. This indicates that large variability in ASD prevalence estimates across studies can be solely due to methodological differences. This is supported by several reviews of ASD epidemiology, pointing out how differences in case ascertainment methods and diagnostic criteria across studies can account for a large proportion of the variance in ASD prevalence estimates; the more thorough examination of the study population and the wider the criteria applied, the higher the prevalence [2]. A recent study also showed that prevalence rates based on clinically diagnosed cases failed to identify many of the research diagnosed cases, thereby leading to lower prevalence estimates [3]. As predicted from this, the DISCO examination generated more cases than the DAWBA, and the prevalence estimated was almost twice as high using the DISCO as when using the DAWBA within the same study (BCS).

Non-response bias

The design of the BCS provides unique opportunities to examine the impact of non-response bias. In spite of the high rate of 70 % participation in the first phase, we saw a clear effect of non-response to the first phase with higher prevalence estimates in the *anonymous ASSQ sample* than in the *identified ASSQ sample* to the extent that overall prevalence estimate rose from 0.72 to 0.87%.

From the first phase to the third phase, although attrition was large, no non-response bias important for the ASD prevalence estimate appears to have been in operation, judging from the scores of the 26% of ASSQ screen-positive children from the first phase who participated in the third phase. However, when we looked at the ASSQ screen-negative children in the third phase, there were significant differences in both parent and teacher ASSQ scores compared to screen-negative children in the first phase. These findings indicate that the ASSQ screen-positive children examined in the third phase were representative of the ASSQ screen-positive children in the first phase with respect to ASD prevalence, whereas this was not true for the ASSQ screen-negative children. The selection bias on behalf of ASSQ screen-negative children was due to the design of the BCS study, as (1) *all* screen-positive children (overall screen) from the first phase were

invited to participate in the second phase, and (2) all children with a DAWBA diagnosis (any disorder) were invited to the third phase. This caused an overrepresentation in the third phase, relative to the first phase, of children who were ASSQ screen-negative but had other mental health problems. Therefore, the ASSQ screen-negative children were excluded from the calculation of the ASD prevalence. True ASD prevalence is therefore probably even higher than 0.87%, as there will always be a few cases among screen-negative children, unless the screening tool used is 100% sensitive. The ASSQ sensitivity was, however, very good [27], and the relative contribution of false ASSQ screen-negative children to overall prevalence of ASD is likely to be low. Previous studies have reasoned likewise [2].

The results of the study indicate that the effects of non-response may vary within a study from phase to phase, and that the relevance of non-response cannot be understood in terms of the magnitude of the non-response, as indicated by Stang [30]. In the first stage non-responders had more difficulties and more high-scorers, to the extent that overall prevalence estimates was affected, even though only 30% were non-responders in the first phase. However, in subsequent phases, the non-response did not influence the prevalence estimate, even though the magnitude of the non-response was up to 74%. Most importantly, our findings indicate that absolute non-responders, i.e. families who do not enrol in a study, are likely to experience higher rate of problems than responders.

Conclusions

A DAWBA diagnosis of ASD is valid, but the DAWBA fails to identify ASD in many higher-functioning cases, thereby underestimating ASD prevalence as compared to the DISCO generated prevalence. The minimum prevalence for all ASD was 0.21%, estimated prevalence 0.87%, but the “true” prevalence may be as high as 1% or above. Our results indicate that there is a correlation between ASD and the propensity for non-response, so that when estimating prevalence, adjustments are necessary in allowing for a higher rate of difficulties in “absolute” non-responders, i.e. children who do not participate even in the questionnaire phases of a study.

Even though we used the same diagnostic criteria, and the study was performed in the same time frame and location, using a more thorough assessment (the DISCO vs. the DAWBA) almost doubled the prevalence estimate within the same population. Such large differences indicate that we need to be cautious when comparing prevalence rates across studies using different methodology, even when studies are using identical diagnostic criteria. It is

important to convey the message that the rise in ASD prevalence could be entirely due to changes in diagnostic criteria, awareness and other, non-aetiological causes. As professionals, we should regard prevalence differences across studies as possibly primarily due to different methodology at various levels, and not give the public unfounded cause for concern about an autism epidemic. When planning epidemiologic studies of ASD one should be aware that the choice of assessment directly affects the results. The expense of in-depth analysis makes it unavailable to the great numbers, but the large error variance and inaccuracy of a questionnaire or semi-structured interview may render it less useful as a measure of clinically important problems [4, 5].

Strengths and limitations

The main strength of the study was having teacher ASSQ from the entire population, making it possible to infer the effects of non-response. In spite of a high participation rate of 70% in the first phase, there was a clear effect of non-response, with higher problem scores and high-scorers among non-responders.

The main limitations are the complicated design of the study with non-random sampling of ASD screen-negative children, the relatively small size of the third phase of the study, and the relatively small number of children diagnosed with ASD.

Acknowledgments The present study was supported by fellowship from the Centre of Child and Adolescent Mental Health, Unifob Health, Bergen, and was also funded by the University of Bergen, the Norwegian Directorate for Health and Social Affairs, the Norwegian Research Council, and the Western Norway Regional Health Authority. We are grateful to the children, parents and teachers for participating in the BCS, and to the other members of the project group for making the study possible. We thank Jim Stevenson and Berit Hilt for helpful comments on the manuscript.

References

1. American Psychiatric Association (2000) Diagnostic and statistical manual of mental disorders, 4th edn (text revision). American Psychiatric Association, Washington
2. Baird G, Simonoff E, Pickles A, Chandler S, Loucas T, Meldrum D et al (2006) Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the special needs and autism project (SNAP). *Lancet* 368(9531):210–215
3. Barbaresi WJ, Colligan RC, Weaver AL, Katusic SK (2008) The incidence of clinically diagnosed versus research-identified autism in Olmsted county, Minnesota, 1976–1997: results from a retrospective, population-based study. *J Autism Dev Disord* Epub ahead of print
4. Brugha TS, Bebbington PE, Jenkins R, Meltzer R, Taub NA, Janas M et al (1999) Cross validation of a general population survey diagnostic interview: a comparison of CIS-R with SCAN ICD-10 diagnostic categories. *Psychol Med* 29(5):1029–1042

5. Brugha TS, Bebbington PE, Jenkins R (1999) A difference that matters: comparisons of structured and semi-structured psychiatric diagnostic interviews in the general population. *Psychol Med* 29(5):1013–1020
6. Fombonne E (2003) Epidemiological surveys of autism and other pervasive developmental disorders: an update. *J Autism Dev Disord* 33(4):365–382
7. Fombonne E, Simmons H, Ford T, Meltzer H, Goodman R (2001) Prevalence of pervasive developmental disorders in the British nationwide survey of child mental health. *J Am Acad Child Adolesc Psychiatry* 40(7):820–827
8. Fombonne E, Zakarian R, Bennett A, Meng L, McLean-Heywood D (2006) Pervasive developmental disorders in Montreal, Quebec, Canada: prevalence and links with immunizations. *Pediatrics* 118(1):e139–e150
9. Gerrits MH, van den Oord EJ, Voogt R (2001) An evaluation of nonresponse bias in peer, self, and teacher ratings of children's psychosocial adjustment. *J Child Psychol Psychiatr* 42(5):593–602
10. Gillberg C, Wing L (1999) Autism: not an extremely rare disorder. *Acta Psychiatr Scand* 99(6):399–406
11. Goodman R (1997) The strengths and difficulties questionnaire: a research note. *J Child Psychol Psychiatr* 38(5):581–586
12. Goodman R, Ford T, Richards H, Gatward R, Meltzer H (2000) The development and well-being assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *J Child Psychol Psychiatr* 41(5):645–655
13. Hansen V, Jacobsen BK, Arnesen E (2001) Prevalence of serious psychiatric morbidity in attenders and nonattenders to a health survey of a general population: the Tromso Health Study. *Am J Epidemiol* 154(10):891–894
14. Haslam N (2006) Bias in psychopathology research. *Curr Opin Psychiatry* 19(6):625–630
15. Heiervang E, Stormark KM, Lundervold AJ, Heimann M, Goodman R, Posserud MB et al (2007) Psychiatric disorders in Norwegian 8- to 10-year-olds: an epidemiological survey of prevalence, risk factors, and service use. *J Am Acad Child Adolesc Psychiatry* 46(4):438–447
16. Honda H, Shimizu Y, Rutter M (2005) No effect of MMR withdrawal on the incidence of autism: a total population study. *J Child Psychol Psychiatr* 46(6):572–579
17. Hysing M, Elgen I, Gillberg C, Lie SA, Lundervold AJ (2007) Chronic physical illness and mental health in children. Results from a large-scale population study. *J Child Psychol Psychiatr* 48(8):785–792
18. Haapea M, Miettunen J, Veijola J, Lauronen E, Tanskanen P, Isohanni M (2007) Non-participation may bias the results of a psychiatric survey: an analysis from the survey including magnetic resonance imaging within the Northern Finland 1966 Birth Cohort. *Soc Psychiatry Psychiatr Epidemiol* 42(5):403–409
19. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P et al (1997) Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry* 36(7):980–988
20. Leekam SR, Libby SJ, Wing L, Gould J, Taylor C (2002) The diagnostic interview for social and communication disorders: algorithms for ICD-10 childhood autism and wing and gould autistic spectrum disorder. *J Child Psychol Psychiatr* 43(3):327–342
21. Lord C, Risi S, Lambrecht L, Cook EH Jr, Leventhal BL, DiLavore PC et al (2000) The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord* 30(3):205–223
22. Lord C, Rutter M, Le Couteur A (1994) Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord* 24(5):659–685
23. Lundberg I, Damstrom Thakker K, Hallstrom T, Forsell Y (2005) Determinants of non-participation, and the effects of non-participation on potential cause-effect relationships, in the PART study on mental disorders. *Soc Psychiatry Psychiatr Epidemiol* 40(6):475–483
24. Madsen KM, Hviid A, Vestergaard M, Schendel D, Wohlfahrt J, Thorsen P et al (2002) A population-based study of measles, mumps, and rubella vaccination and autism. *N Engl J Med* 347(19):1477–1482
25. Mullick MS, Goodman R (2005) The prevalence of psychiatric disorders among 5–10 year olds in rural, urban and slum areas in Bangladesh: an exploratory study. *Soc Psychiatry Psychiatr Epidemiol* 40(8):663–671
26. Posserud MB, Lundervold AJ, Gillberg C (2006) Autistic features in a total population of 7–9-year-old children assessed by the ASSQ (autism spectrum screening questionnaire). *J Child Psychol Psychiatr* 47(2):167–175
27. Posserud M, Lundervold AJ, Gillberg C (2008) Validation of the autism spectrum screening questionnaire in a total population sample. *J Autism Dev Disord* Epub ahead of print
28. Rutter M, Cox A, Tupling C, Berger M, Yule W (1975) Attainment and adjustment in two geographical areas. I—the prevalence of psychiatric disorder. *B J Psychiatry* 126:493–509
29. Sponheim E, Skjeldal O (1998) Autism and related disorders: epidemiological findings in a Norwegian study using ICD-10 diagnostic criteria. *J Autism Dev Disord* 28(3):217–227
30. Stang A (2003) Nonresponse research—an underdeveloped field in epidemiology. *Eur J Epidemiol* 18(10):929–931
31. Stormark KM, Heiervang E, Heimann M, Lundervold A, Gillberg C (2008) Predicting nonresponse bias from teacher ratings of mental health problems in primary school children. *J Abnorm Child Psychol* 36(3):411–419
32. Volkmar FR, Lord C, Bailey A, Schultz RT, Klin A (2004) Autism and pervasive developmental disorders. *J Child Psychol Psychiatr* 45(1):135–170
33. Wakefield AJ, Murch SH, Anthony A, Linnell J, Casson DM, Malik M et al (1998) Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet* 351(9103):637–641
34. Williams JG, Higgins JP, Brayne CE (2006) Systematic review of prevalence studies of autism spectrum disorders. *Arch Dis Child* 91(1):8–15
35. Wing L, Leekam SR, Libby SJ, Gould J, Larcombe M (2002) The diagnostic interview for social and communication disorders: background, inter-rater reliability and clinical use. *J Child Psychol Psychiatr* 43(3):307–325
36. Wing L, Potter D (2002) The epidemiology of autistic spectrum disorders: is the prevalence rising? *Ment Retard Dev Disabil Rev* 8(3):151–161