



RESEARCH ARTICLE

Validating the Developmental and Well-Being Assessment (DAWBA) in a clinical population with high-functioning autism [version 1; peer review: awaiting peer review]

Nadia Coscini ¹, Ramya Srinivasan ², David Skuse²

¹General Medicine, Royal Children's Hospital, Melbourne, Australia

²Population, Policy & Practice Department, UCL Great Ormond Street Institute of Child Health, London, UK

v1 **First published:** 17 Jun 2020, 9:622
<https://doi.org/10.12688/f1000research.24409.1>
Latest published: 17 Jun 2020, 9:622
<https://doi.org/10.12688/f1000research.24409.1>

Abstract

Background: With increasing numbers of referrals to health services for assessment of Autism Spectrum Disorder (ASD), the Developmental and Well-Being Assessment (DAWBA) has been suggested as a useful screening instrument to assist in prioritising patients for review. It is an online interview for parents that has been previously validated for ASD in a non-clinical community sample of twins. Our study aimed to evaluate its predictive validity in a complex clinically-referred sample of children with suspected high-functioning autism.

Methods: The sample comprised 136 children (females = 53; males = 83) who were referred for ASD assessment at the Social Communication Disorder Clinic (SCDC) at Great Ormond Street Hospital. Parents completed the DAWBA online prior to undergoing a multi-disciplinary team (MDT) assessment. This included completing the Developmental, Dimensional and Diagnostic Interview (3di) and the Autism Diagnostic Observation Schedule (ADOS). Two clinicians independently rated the DAWBA using DSM-5 diagnostic criteria and compared results to the MDT outcome, which was considered gold standard.

Results: Compared with an MDT assessment, the DAWBA interview demonstrated good sensitivity (0.91) but poor specificity (0.12). Overall, 64% of cases were accurately assigned as case/non-case. Estimates of positive (0.66) and negative (0.43) predictive validity were influenced by the relatively high prevalence of ASD in the study sample (65%).

Conclusion: The DAWBA online interview has excellent sensitivity in a clinical population of complex neurodevelopmental disorders, containing a high prevalence of ASD, but specificity was poor. As the SCDC offers tertiary opinions on disputed cases of suspected ASD, the population cohort limits the generalisability of these results. Further evaluation is required in community child mental health or paediatric services.

Keywords

Autism spectrum disorder, assessment, validity, sensitivity and specificity, social communication disorder, screening instrument.

Open Peer Review

Reviewer Status *AWAITING PEER REVIEW*

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the **UCL Child Health**



gateway.

Corresponding author: Nadia Coscini (nadia.coscini@gmail.com)

Author roles: **Coscini N:** Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Srinivasan R:** Data Curation, Formal Analysis, Methodology, Project Administration, Supervision, Validation, Writing – Review & Editing; **Skuse D:** Conceptualization, Data Curation, Supervision, Validation, Visualization, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by a Royal Children's Hospital Foundation Travelling Grant.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2020 Coscini N *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Coscini N, Srinivasan R and Skuse D. **Validating the Developmental and Well-Being Assessment (DAWBA) in a clinical population with high-functioning autism [version 1; peer review: awaiting peer review]** F1000Research 2020, 9:622 <https://doi.org/10.12688/f1000research.24409.1>

First published: 17 Jun 2020, 9:622 <https://doi.org/10.12688/f1000research.24409.1>

Introduction

Autism Spectrum Disorder

The Diagnostic and Statistical Manual (5th edition) describes Autism Spectrum Disorder (ASD) as a condition with social communication problems and patterns of repetitive behaviour or restricted interests (American Psychiatric Association, 2013). The prevalence of autism in England is 1.3% based on figures from the United Kingdom Department for Education (Department for Education (UK), 2017). There have been increasing numbers of referrals for autism, with services receiving twice the previous number, but without the required resources with which to make assessments (Dreaper, 2017). At an estimated cost of £800 per child per diagnostic assessment, service providers are under increasing pressure to assess and prioritise patients (Galliver *et al.*, 2017). With recognition of the benefits of earlier diagnosis, there has also been a focus on earlier recognition of ASD. A recent study has shown that in the UK, the median age for being diagnosed with ASD is 4.5 years, but despite strategies to try and reduce this, it has remained the same between 2004 and 2014 (Brett *et al.*, 2016). Therefore, there may be a benefit in screening potential cases of ASD in order to triage the need for further assessment. Application of a diagnostic tool prior to a multidisciplinary assessment in order to obtain a clinical history could also inform the diagnostic process and expedite the clinical evaluation.

Comorbid disorders

There is a high prevalence of comorbid psychiatric disorders in children with ASD, with one study estimating this as up to 70% (Simonoff *et al.*, 2008). Whilst there is a focus on earlier diagnosis of ASD, there is no corresponding emphasis on the diagnosis of comorbid disorders (Mannion & Leader, 2013). It is particularly important to recognise comorbidities because of the increased impact on functioning and quality of life for children diagnosed with ASD (Soke *et al.*, 2018). Understanding comorbid disorders also has the potential to make the ASD diagnosis more accurate.

The diagnostic process

The best practice guidelines for diagnosing ASD recommend that a multi-disciplinary team (MDT) should undertake a thorough medical and developmental history, clinical examination and a semi-structured period of observation (NICE, 2011). The guidelines recommend the MDT should involve a paediatrician or psychiatrist, psychologist and speech pathologist and also suggest an autism assessment is helpful, but do not recommend any specific one (NICE, 2011). Frequently used tools include the Developmental, Dimensional and Diagnostic Interview (3di) (Skuse *et al.*, 2004), the Autism Diagnostic Interview-Revised (ADI-R) (Lord *et al.*, 1994a), and the Autism Diagnosis Observation Schedule (ADOS) (Falkmer *et al.*, 2013). These interviews can take between 1.5–3 hours to complete and can only be implemented with trained personnel (McEwan & Brinkmann, 2000). Due to the length of these assessments, they are useful as part of the MDT but not as potential screening instruments to determine which children are most at risk and should be prioritised for review.

The need for a screening tool or diagnostic instrument

Before evaluating whether a screening tool may also have utility as a diagnostic instrument, it is useful to review its nature and function. The United Kingdom National Screening Committee defines it as a test that is applied in a systematic way to a specific population to identify people at risk in order to investigate further or prevent the disease in an individual who would not otherwise have looked for the disorder (The UK National Screening Committee, 1998, in Gilbert *et al.*, 2001). In this situation, screening would not be for all children in the general population, but only for those about whom concerns have been raised about possible developmental problems, who are then subsequently referred for paediatric or psychiatric assessment. Fundamental values of a screening test were introduced by Wilson and Jungner in their seminal 1968 paper, which highlighted that it should be cheap, simple to administer, acceptable, reliable, valid and be able to be followed up appropriately (Wilson & Jungner, 1968). There is a need for an assessment that is inexpensive to administer prior to being seen in clinic that can help clinicians identify individuals most at risk of ASD.

Since a screening test of ASD is likely to collect a large amount of data, this could also be utilised to its maximum capacity by integrating the results into a subsequent multi-disciplinary assessment. This would mean that a screening test could also be used as a diagnostic adjunct or instrument providing information to clinicians to assist them reaching a conclusion about whether an individual has autism. It may improve the efficiency of diagnosis by speeding up information-gathering since much of the relevant information would be collected prior to the diagnostic process. It may also support decision-making in difficult or unclear cases, and assist in gathering broader information about emotional and behavioural problems, helping to clarify the diagnosis when a specific presentation may be caused by different conditions.

The Developmental and Well-Being Assessment (DAWBA) may be useful in both these ways and its development was envisioned as such by its authors (Goodman *et al.*, 2000). It has already been used in epidemiological mental health surveys in the UK for young people (Meltzer *et al.*, 2000).

The Developmental and Wellbeing Assessment

The Developmental and Well-Being Assessment (DAWBA) is an instrument for caregivers to complete that helps identify possible psychiatric disorders (Goodman *et al.*, 2000). It is an interview that is completed online, which has been used in many research studies and all national UK studies looking at psychopathology in children and adolescents. The information gathered helps guide clinicians on areas of risk before the child attends clinic. It can be completed by multiple different informants including children over 11 years of age, their parents or main caregivers, and school teachers. The DAWBA has an incorporated diagnostic algorithm that can predict risk of psychiatric morbidity, and clinical raters can analyse the unstructured answers and synthesise information from the structured questions across the different disorder modules to refine the DAWBA-generated diagnoses.

Any screening or diagnostic tool should be able to be applied to cases that are difficult to diagnose in ASD, due to the spectrum of symptoms or the presence of other comorbidities. There are other questionnaires used by some services as a screening tool, but they focus only on ASD symptoms. The advantage of using the DAWBA, is that it covers a wide range of psychopathologies and can be completed online in the parents' own time. Furthermore, its need for review by a clinician means that it is not just a questionnaire relying on a statistical cut off, and context can be applied to interpreting the symptoms reported.

The DAWBA's validity

The initial validity study for the DAWBA was performed in both clinical and community samples. In the clinical sample of patients in a child and adolescent mental health clinic, there was strong agreement between the primary diagnosis yielded by the DAWBA and that of the clinical case notes. There was however, an increase in comorbid diagnoses made that were either only possible (not definitive) or completely absent in the case notes (Goodman *et al.*, 2000). Within the community sample, the respondents reflected the expected diversity of clinical characteristics, and the DAWBA also found the expected difference in rates of diagnosis between a clinical and community sample (Goodman *et al.*, 2000). There have been further studies that have validated the DAWBA, and it has also been used and validated in epidemiological studies (Ford *et al.*, 2003; Heiervang *et al.*, 2007). Two studies found high values of agreement between the DAWBA and clinical diagnosis (Alyahri & Goodman, 2006; Mullick & Goodman, 2005). Alyahri and Goodman compared the DAWBA and Strengths and Difficulties Questionnaire (SDQ) in a Yemeni population, including children in community psychiatric clinics, and found good agreement for internalising and externalising disorders (Alyahri & Goodman, 2006). The other validation study used the Bangladeshi translation of the DAWBA in a cohort of 100 children who had mental health referrals. The DAWBA diagnoses demonstrated a high value of agreement compared with clinical diagnoses (Mullick & Goodman, 2005). Another study looked at how the DAWBA performed when used as an adjunct to clinical diagnosis in a Swiss population of children referred to an outpatient mental health service (Aebi *et al.*, 2012). Of 270 consecutive new referrals, half of the clinicians were randomised to receive data from the DAWBA as well as the clinical rater diagnosis, while the other half did not (Aebi *et al.*, 2012). Overall there was fair to moderate agreement for all diagnoses, whether clinicians received data or not.

Throughout these studies, inter-rater reliability between clinical raters has been reported to be consistently high (Aebi *et al.*, 2012; Fleitlich-Bilyk & Goodman, 2004; Ford *et al.*, 2003; Heiervang *et al.*, 2007).

Validity of the DAWBA for ASD

In a 2016 study looking at participants from a national twin study, the DAWBA was used to assess possible ASD and demonstrated high sensitivity (0.86) and specificity (0.87) (McEwen *et al.*, 2016). This large population of twins was screened for ASD and parents completed the DAWBA interview. A small proportion of this group were then allocated into

at-risk groups for ASD based on their DAWBA scores; high risk or low risk. Within these groups of predicted ASD and non-ASD, a subsample was followed up by an at-home visit by researchers completing a gold standard MDT ASD assessment, including the ADOS. As the sample in this study was community-based, and assessed based on at-risk screening, the utility for the DAWBA in a clinical setting needs to be tested. Its ability to predict an ASD diagnosis needs to be validated in a clinically-referred sample.

Aims and objectives

We propose comparing the outcomes of DAWBA- and MDT-diagnostic assessments for children who have been referred to a developmental clinic for concerns about ASD. If the DAWBA proves to have good predictive validity, then it may be useful as an efficient screening measure not only for ASD, but also for comorbid psychiatric disorders. This would allow clinicians and researchers to utilise the DAWBA with confidence. We expect, given the excellent history of inter-rater reliability, that the DAWBA will have good inter-rater reliability between clinical raters. Further to McEwan's study, we are keen to find out how sensitive and specific the DAWBA is in diagnosing ASD in a clinically complex referral group which in this study was the Social and Communication Disorder Clinic, which is discussed below. The DAWBA may be useful as a screening test before clinical assessment to highlight a possible risk of diagnosis, which could be used as part of diagnosing ASD. If the DAWBA is validated for ASD, it is also possible that service providers could utilise the interview as part of a triaging system to prioritise referrals for assessment.

Methods

Participants and procedures

The Social Communication Disorder Clinic (SCDC) and the multi-disciplinary team (MDT) assessment. The sample in this study included all individuals aged 4–18 years who were referred to the Social Communication Disorder Clinic (SCDC) at Great Ormond Street Hospital, and completed a DAWBA before a gold standard MDT ASD diagnostic assessment (n=136) between June 2014 and August 2017. Parents of children seen in the SCDC were asked to whether they were interested in participating in the Autism Families Study, using data collected for future research on autism spectrum disorders (Great Ormond Street Hospital for Children NHS, 2020).

The SCDC is a tertiary specialist clinic, which focusses on providing an extensive and evidence-based multi-disciplinary assessment for children referred from local community clinics who are uncertain about a child's diagnosis of possible ASD. The SCDC implements the NICE guidelines for a gold standard MDT assessment for ASD, by incorporating a thorough patient medical history and examination, followed by a structured developmental interview (the 3di) and standardised observation (the ADOS). All prior assessments and reports are collated and reviewed by the MDT, and members of the team may also observe the child in school as part of the wrap-around approach to assess the child in all their environments. Parents of children referred to the clinic were asked to complete the DAWBA before their first appointment at the SCDC.

As a tertiary referral clinic, the children referred to the SCDC are not a typical population of children referred for assessment of autism in general. All the children were referred for second opinions on previous autism assessments. Many of the children had been referred because health professionals were unsure about whether they met the full criteria for ASD due to a complex clinical presentation.

All cases were reviewed and data collected on the results of ADOS scores (overall and modules), 3di subscale scores, Strengths and Difficulties Questionnaire (SDQ) scores, DAWBA computer predicted probabilities, DAWBA consensus ratings, and final MDT diagnosis. We also collected demographic data including parental level of education and pathway of referral. Pathways to referral were divided into two categories: via a paediatrician or by local Child and Adolescent Mental Health Services (CAMHS). Information about parental education was categorised both by whether parents had completed high school or not (defined as A levels at 17–18 years of age in this United Kingdom cohort).

Confidentiality and consent. The standard protocol for confidentiality was followed according to the Data Protection Act 1998 (DPA) adopted by University College London and Great Ormond Street Hospital. Each child seen in clinic has details stored in a secure database that is only available to staff involved either with the clinic or approved research. Each DAWBA record has a unique de-identified number and records are not identifiable based on patient information. All families who were reviewed in the SCDC provided written informed consent to participate in future research through the Autism Families Study (Great Ormond Street Hospital for Children NHS, 2020). Ethics approval was already in place as per the Autism Families Study (REC reference number is 06/Q0508/60 and R&D number is 08BS06).

Tool to be validated: The DAWBA

The DAWBA was completed by parents online or by trained research students who conducted the interview via phone and completed the online form. The DAWBA consists of a computer program using a structured interview format with extended answers that then uses an algorithm to predict psychiatric diagnoses. The first part of the DAWBA is the Strengths and Difficulties Questionnaire, which is further discussed in detail below. Respondents are then asked about symptoms of up to seventeen types of psychiatric diagnoses, including ASD. With these answers, the DAWBA's algorithm uses its calculated scores to generate a probability of diagnosis for each disorder. These computer-predicted probabilities are termed the "DAWBA bands".

DAWBA 'skip rules' were employed in order to gather more focused data in areas where the respondent gives a set number of positive responses. Respondents also provided full text answers to open-ended questions, providing further description to assist clinical raters in their interpretation.

The Strengths and Difficulties Questionnaire. An overall understanding of the social and emotional wellbeing of each

subject was measured using the Strengths and Difficulties Questionnaire (SDQ) section of the DAWBA. The SDQ is an interview that measures social and emotional difficulties by assessing the severity and impact of problems in five domains listed in order of assessment: emotional problems, conduct issues, hyperactivity, problems with peers, and prosocial behaviour. There are 20 questions on difficulties, and 5 questions on prosocial behaviour. Total difficulty scores were calculated by adding up the scores on the 20 items relating to difficulties. Respondents were able to select from 3 possibilities ranging from: "certainly true", "somewhat true" and "not true".

Total scores were reported in comparison to population norms generated by a survey of mental health in 5–15-year-old children in the United Kingdom (Meltzer *et al.*, 2003). These overall reports were classified as increasing in severity from "close to average", "slightly raised", "high" and "very high" for domains that looked at difficulties. Reverse order classification was used for prosocial problems with categories referred to as "close to average", "slightly lowered", "low" and "very low".

The DAWBA predictability bands. The algorithm within the DAWBA analysis tool was used to provide probabilities of various conditions including ASD. The probability was presented as one of six levels (numbered 0–5). Level 0 corresponds to less than 0.1% probability of the subject having the disorder. This continues with level 1 representing approximately 0.5% probability, level 2 being 3% probability, level 3 being 15%, level 4 being 50%, and level 5, greater than 70%.

The DAWBA clinical-rater consensus diagnosis. Using both the DAWBA-generated predictions and free text information provided by responders as part of their decision-making process, two independent clinical raters assessed the records for either a positive or negative ASD diagnosis. Diagnosis was informed by criteria as set out by the Diagnostic and Statistical Manual of Mental Disorders (5th edition; (American Psychiatric Association, 2013)). One assessor was a child psychiatrist who had previous experience as a DAWBA clinical rater. The second assessor was a paediatric clinician. Both the raters only had access to de-identified DAWBA records and were blind to the outcome of the MDT assessment. The clinical raters then met to discuss and compare the results and reached a Consensus diagnosis if there were any disagreements. In 15 cases where there was uncertainty about the Consensus diagnosis, a third senior clinical rater reviewed the cases to make a final diagnosis. The final DAWBA clinical rater diagnoses will be referred to as the Consensus diagnosis throughout the rest of this document.

Measures in the MDT diagnosis

As part of the gold standard MDT assessment completed by the SCDC, the cohort was assessed using the Developmental, Dimensional and Diagnostic Interview (3di) and the Autism Diagnostic Observation Schedule (ADOS).

The Developmental, Dimensional and Diagnostic Interview (3di). The 3di was administered by trained psychologists working as part of the SCDC team. The 3di is an interview based on parent-report which measures symptoms of ASD and associated

comorbidities (Skuse *et al.*, 2004). The 3di asks questions about the child's sociodemographic information, developmental history and motor coordination which is entered into a computer program. Once the semi-structured interview is completed, the computer algorithm is able to immediately generate a short report. In total, there are 183 questions on sociodemographic details, motor and development history, 266 questions associated with ASD and 291 questions based on other mental disorders. Scores are generated based on the responses: 0 is given for no record of any symptoms, 1 is for small amount of evidence of asked-about behaviour, and 2 is for known or ongoing presence of the behaviour. Once the interviews were completed, the final scores were divided into three components representative of ASD: Reciprocal social interaction, Communication, and Repetitive and stereotyped patterns of behaviour. These scores were then used as equivalent ADI-Algorithm scores as discussed in the next section.

The 3di scores and ADI-Algorithm. The Autism Diagnostic Interview, Revised (ADI-R) is an interview that assesses the three domains impaired by ASD (social reciprocity, communication and restricted behaviours) and makes a diagnosis based on these three domains. The 3di was also modelled on these three domains and therefore subscale scores from the 3di can be used to calculate scores which are equivalent to the ADI-R (Santosh *et al.*, 2009). When these scores are derived from the 3di they are termed the ADI-Algorithm scores.

The 3di conducted at the SCDC produces scores that reflect possible impairment in one of the following domains; Social reciprocity, Communication and Restricted behaviours. As the 3di output score is consistent with an ADI-Algorithm score, all results from the 3di are henceforth referred to as the ADI-Algorithm. The ADI-Algorithm scores were analysed based on the three separate domains, and they were also grouped together to look at Social reciprocity and Communication and all three combined.

The Autism Diagnostic Observation Schedule (ADOS). The ADOS was administered by trained psychologists and allied health specialists working as part of the SCDC team. The ADOS is a semi-structured instrument that measures the observed behaviour of children for possible ASD, and two currently available and validated versions were used (ADOS and ADOS-2). The assessment uses four modules that look at the domains of play, imagination, social interaction and communication. There are four modules that can be used based on a child's level and each takes approximately 30 minutes to complete (Lord *et al.*, 2000). There is a total score at the completion of the modules and these are based on the individual scores of the social and communication domains, and the sum of the play and imagination domains (Oosterling *et al.*, 2010). We identified a positive result on the ADOS if it diagnosed Autism. If the threshold showed a category of Autism Spectrum or no diagnostic features, we classified it as a negative result.

Data analysis

Analysis was completed using data stored on Microsoft Excel and analysed with IBM SPSS Statistics 25.0 (IBM Corp, 2017).

We investigated differences between groups using chi-square tests or two-sample *t*-tests and set levels of significance at 0.05. Characteristics of the sample considered the background demographics of the participants by generating a table looking at the number of children who received a diagnosis of ASD, their gender, level of maternal and paternal education, the referral pathway and the child's age in years.

The Cohen's kappa statistic was used to calculate the inter-rater reliability of the independent DAWBA clinical raters. The percentage of agreement was also measured for comparison.

The mean SDQ scores of the ASD and non-ASD groups were compared. These scores were then compared to mean scores in the general population. The mean score results for the ADOS and ADI-Algorithm were also compared between ASD and non-ASD groups. The ADI-Algorithm mean scores were divided into: Social reciprocity; Communication; Combined Social reciprocity and Communication scores; Restricted behaviours; and, Combined Social reciprocity, Communication and Restricted behaviours scores. Comparison between all the scores listed in the SDQ, ADOS and ADI-Algorithm were checked for significance using 2-sample *t*-tests.

The sensitivity, specificity and positive and negative predictive values of the Consensus diagnosis compared to the MDT diagnosis were calculated. The diagnostic results from the ADOS and ADI-Algorithm were also compared to the MDT diagnoses. We further compared the result of the Consensus diagnosis directly with the diagnostic results of the ADOS and ADI-Algorithm.

Finally, we were interested to examine the predictive validity of the DAWBA algorithm relative to the Consensus diagnoses. Therefore, we calculated the sensitivity, specificity, PPV and NPV of five predictability bands of the DAWBA for ASD diagnosis compared with the MDT results. We measured the accuracy of the DAWBA probability bands in predicting ASD against the MDT diagnosis using a receiver operating characteristic (ROC) curve.

Results

Sample characteristics

A total of 149 participants completed the DAWBA before attending the Social Communication Disorder Clinic at GOSH. Of those that completed the DAWBA and underwent a multidisciplinary team assessment, 13 were excluded due to an incomplete DAWBA. This is demonstrated in Figure 1.

A final total of 136 participant records were analysed. Of the total participants, 64% (87) were diagnosed by the MDT as having ASD. The mean age of the cohort was similar in both subgroups; 10.5 years for those with ASD compared to 10.4 for the non-ASD group. There was a similar rate of males diagnosed with ASD (59%) compared to those who were not diagnosed (65%). Table 1 lists the sample characteristics of the cohort. There was no evidence of any difference between the groups.

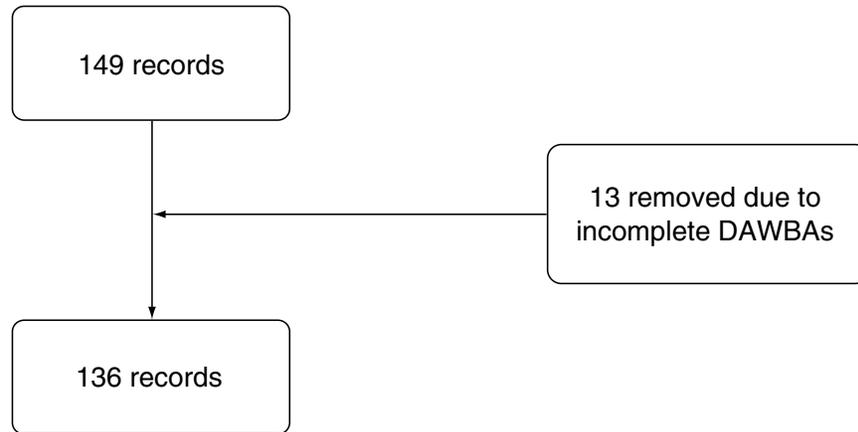


Figure 1. Flowchart of participants.

Table 1. Sample characteristics.

	ASD n (%)	Non-ASD n (%)	p value for difference
Child gender, male	51 (59)	32 (65)	0.47
Maternal education, completed high school	55 (74)	36 (86)	0.17
Paternal education, completed high school	45 (71)	29 (83)	0.23
Referred by paediatrician	56 (64)	32 (67)	0.85
	M (SD)	M (SD)	
Child age in years	10.5 (3.4)	10.4 (3.0)	0.83
Total number of children	87	49	

ASD, Autism Spectrum Disorder.

Parents of non-ASD children had a trend towards slightly higher level of education compared to parents of children diagnosed with ASD. 85% of those without ASD had a mother who completed higher education compared to 75% those with ASD. Similarly, 82% of those assessed as non-ASD had fathers who completed A levels compared to 72% who were diagnosed with ASD. The referral pathway for both groups was also quite similar. 63% of children with ASD were referred to the SCDC by their paediatrician compared to 67% in the non-ASD group. The rest were referred by a Child and Adolescent Mental Health Services (CAMHS) clinician.

The Strengths and Difficulties Questionnaire

The SDQ mean scores demonstrate the difference between the populations. As seen in Table 2, the mean scores for each section between the ASD and non-ASD groups are similar. There was evidence for a difference in degree of conduct problems (CP) between those who were diagnosed with ASD and those who were not (Mean CP subscale score – ASD: 3.8, SD 2.0, 95% CI 3.3–4.2; Non-ASD: 5.0, SD 2.6, 95% CI 4.3–5.8; p = <0.01). There was no evidence of difference on any of the other SDQ subscales between the two groups.

SDQ scores can be represented categorically based on previously published cut-offs (Goodman, 2018). Table 3 visually represents the categorical values of the mean scores in both groups compared to the population norm.

ADOS mean scores

Table 4 demonstrates the mean score for the ADOS between the ASD and non-ASD groups. There was strong evidence for a difference in the total ADOS score between those who were diagnosed with ASD and those who were not (Mean ADOS score – ASD: 12.1, SD 6.2, 95% CI 10.8–13.5; Non-ASD: 6.0, SD 4.5, 95% CI 4.5–7.4; p= <0.01).

ADI-Algorithm mean scores

Table 5 demonstrates the mean scores for the ADI-Algorithm between the ASD and non-ASD groups. There was strong evidence for a difference in the Restricted behaviours scale score between those who were diagnosed with ASD and those who were not (Mean Restricted behaviours score – ASD: 4.8, SD 2.4, 95% CI 4.3–5.3; Non-ASD: 3.5, SD 2.1, 95% CI 2.9–4.2 p= 0.003). There was no evidence of difference on the other domain scores between the two groups.

Table 2. Mean scores of the Strengths and Difficulties Questionnaire.

Domain	Scores				p value ASD vs Non-ASD	UK National Norm
	ASD		Non-ASD			
	M (SD)	95% CI	M (SD)	95% CI		
Total difficulties	22.5 (6.5)	21.2–23.9	23.7 (6.1)	21.9–25.5	0.28	8.4 (5.8)
Emotional symptoms	6.0 (2.8)	5.4–6.6	6.3 (2.7)	5.5–7.0	0.63	1.9 (2.0)
Conduct problems	3.8 (2.0)	3.3–4.2	5.0 (2.6)	4.3–5.8	<0.01*	1.6 (1.7)
Hyperactivity	7.4 (2.3)	7.0–7.9	7.9 (2.1)	7.2–8.5	0.31	3.5 (2.6)
Peer problems	5.3 (2.5)	4.8–5.8	4.5 (2.5)	3.8–5.3	0.14	1.5 (1.7)
Prosocial	5.0 (2.6)	4.4–5.5	4.7 (2.1)	4.1–5.3	0.60	8.6 (1.6)

Norms provided by (Meltzer *et al.*, 2000). ASD, Autism Spectrum Disorder.

Table 3. Strengths and Difficulties Categories.

Domain	Category		
	ASD	Non-ASD	Population norm
Total difficulties	Very high	Very high	Close to average
Emotional symptoms	High	High	Close to average
Conduct problems	Slightly raised	High	Close to average
Hyperactivity	High	High	Close to average
Peer problems	Very high	High	Close to average
Prosocial	Very low	Very low	Close to average

Blue, close to average; yellow, slightly raised/slightly lowered; orange, high/low; red, very high/very low. ASD, Autism Spectrum Disorder.

Table 4. Statistics of ADOS in the Autism Spectrum Disorder (ASD) and non-ASD groups.

Score	ASD		Non-ASD		p value for difference
	M (SD)	95% CI	M (SD)	95% CI	
Total ADOS Score	12.1 (6.2)	10.8–13.5	6.0 (4.5)	4.5–7.4	<0.01*

Table 5. Statistics for ADI-Algorithm in the Autism Spectrum Disorder (ASD) and non-ASD groups.

Domain Score	ASD		Non-ASD		p value for difference
	M (SD)	95% CI	M (SD)	95% CI	
Social reciprocity	14.0 (6.4)	12.6–15.4	12.0 (6.3)	10.1–13.8	0.48
Communication	11.7 (6.2)	10.4–13.0	12.3 (4.4)	11.0–13.6	0.66
Combined Social Reciprocity and Communication scores	25.7 (10.4)	23.5–27.9	24.1 (8.8)	21.5–26.7	0.71
Restricted behaviours	4.8 (2.4)	4.3–5.3	3.5 (2.1)	2.9–4.2	<0.01*
Total ADI-Algorithm Score (Combined Social reciprocity, Communication and Restricted behaviours)	30.5 (11.2)	28.1–32.9	27.6 (9.9)	24.6–30.1	0.11

Results of DAWBA Consensus rating

Out of 136 clinical ratings of the DAWBA, 119 (87.5%) had initial agreement. A further two reached Consensus between the two clinical raters. The remaining 15 reached Consensus together with a third clinical rater.

Level of agreement and inter-rater reliability

As discussed above, there was an 87.5% rate of agreement between diagnoses. The inter-rater reliability for the DAWBAs assessed by clinical raters in the SCDC cohort was poor (Cohen’s kappa = 0.43). Before completing the SCDC clinical-rating of the DAWBAs, the authors independently reviewed 308 DAWBA records of a different clinical group with excellent results for inter-rater reliability for ASD (Cohen’s kappa = 0.83) (IMAGINE ID Study, 2018).

Validity results for DAWBA and other tests compared to MDT diagnoses

When comparing the Consensus diagnoses with the MDT diagnoses, the DAWBA showed a sensitivity of 0.91 and a specificity of 0.12. The Positive Predictive Value (PPV) was 0.66 while the Negative Predictive Value (NPV) of the Consensus diagnoses was 0.43 (Table 6).

Table 7 also shows the results for the face-to-face assessments. This result for the ADOS assessment showed a sensitivity of 0.67 and specificity of 0.78, while the PPV was 0.86 and the

NPV 0.53. The ADI-Algorithm showed a sensitivity of 0.53 and specificity of 0.54 when compared with the MDT diagnoses. PPV was 0.69 and NPV 0.38. We also compared it with where just the Social reciprocity and Communication scales met ADI-Algorithm criteria for ASD, which showed a sensitivity of 0.34 and a specificity of 0.63. The PPV was 0.63 and NPV 0.34.

Comparing validity for Consensus diagnosis with ADOS and ADI-Algorithm

We then assessed the validity of the Consensus diagnosis when using the ADOS and ADI-R assessments as the reference standard for diagnosis. The results show a similar pattern to when Consensus is compared to MDT diagnosis with sensitivity and low specificity as per Table 8.

Table 6. Results of DAWBA Consensus rating compared to multi-disciplinary team (MDT) diagnosis.

Consensus diagnosis	MDT diagnosis		
	ASD (n)	Non-ASD (n)	Total
ASD (n)	79	43	122
No ASD (n)	8	6	14
Total	87	49	136

ASD, Autism Spectrum Disorder.

Table 7. Validity results for Consensus diagnosis compared to multi-disciplinary team (MDT) diagnosis, ADOS and ADI-Algorithm diagnostic results.

<i>DAWBA results compared to MDT Diagnosis</i>					
DAWBA result	Sensitivity	Specificity	PPV	NPV	Correctly classified %
DAWBA Consensus rating	0.91	0.12	0.66	0.43	63
<i>Autism interview results compared to MDT Diagnosis</i>					
ADOS	0.67	0.78	0.86	0.53	70
Meets ADI-Algorithm criteria on Social reciprocity and Communication scales	0.34	0.63	0.63	0.34	44
Meets ADI-Algorithm criteria on Social reciprocity, Communication and Repetitive behaviour scales	0.53	0.54	0.69	0.38	54

Table 8. Consensus Diagnosis compared with ADOS and ADI-Algorithm diagnostic results.

Test result	Sensitivity	Specificity	PPV	NPV	Correctly classified %
ADOS	0.94	0.15	0.54	0.96	55
Meets ADI-Algorithm criteria on Social reciprocity and Communication scales	0.93	0.12	0.52	0.62	53
Meets ADI-Algorithm criteria on Social reciprocity, Communication and Repetitive behaviour scales	0.85	0.07	0.33	0.46	34

Comparing Validity for DAWBA prediction bands compared to MDT diagnosis

Table 9 demonstrates the sensitivity and specificity of the DAWBA prediction bands for an ASD diagnosis relative to MDT diagnosis. There were five possible cut-off levels for a positive result. As expected, sensitivity increased as the cut-off was reduced but was never higher than 0.76. The specificity decreased from 0.94 at the highest cut-off to 0.29 at the lowest. Figure 2 demonstrates the distribution of ASD and non-ASD cases within each prediction band.

ROC curve – MDT diagnosis vs DAWBA probability bands

An ROC curve was generated to compare the DAWBA probability bands with the MDT ASD diagnosis. Figure 3 demonstrates that the area under the curve was not significant (AUC = 0.58, 95% CI = 0.5–0.7).

Discussion

Excellent sensitivity, poor specificity

This study found that the DAWBA Consensus diagnosis was accurate in 64% of diagnoses compared to the gold standard MDT. It had good sensitivity (91%) but poor specificity (12%). These findings were unexpected. Based on previous studies, we had anticipated that there would be high sensitivity and specificity in diagnosing ASD. We suspect that the difference between the previous studies and these results relates to the unique quality of our sample population. Due to the DAWBA’s previously documented ability to differentiate autism from non-autism in a community twin sample, we expected that it would perform similarly well in discriminating autism from non-autism in a more complex clinical group referred for assessment. Unfortunately, the complexity of our sample has resulted in the DAWBA performing with poorer specificity

Table 9. Classification statistics for DAWBA prediction bands compared to multi-disciplinary team (MDT) diagnosis.

Computer-generated predictability bands								
Positive when using this cut-off	Number in band	Cumulative Frequency	With ASD*	Sn	Sp	PPV	NPV	% Correctly classified
High (70%+)	14	14	11	0.13	0.94	0.79	0.38	42
Moderate (~50%)	20	34	26	0.30	0.84	0.76	0.40	49
Low (~15%)	28	62	44	0.51	0.63	0.71	0.42	55
Low (~3%)	39	101	66	0.76	0.29	0.65	0.40	59
Low (0.5%)	0	101	66	-	-	-	-	-
Very low (<0.1%)	35	136	87	1.00	0.00	0.64	0.00	64

* according to MDT diagnosis. ASD, Autism Spectrum Disorder.

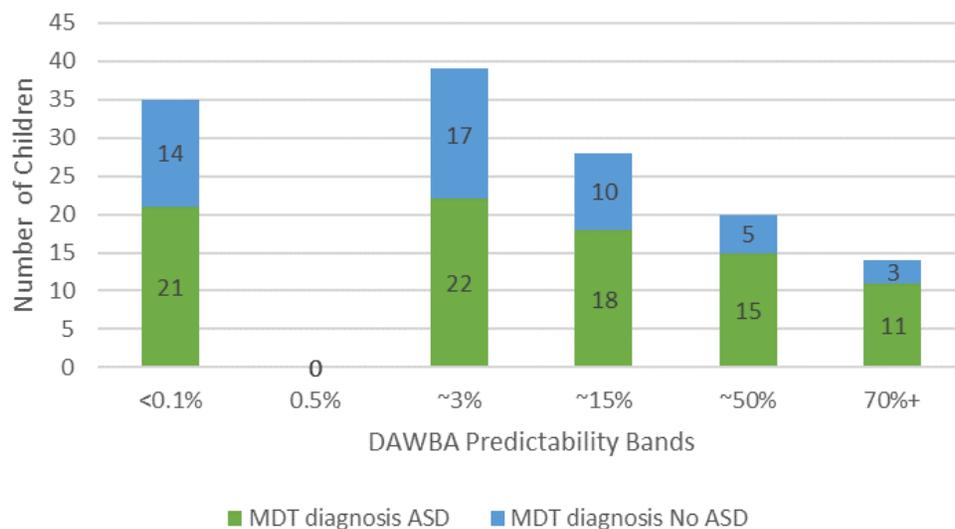


Figure 2. Distribution of DAWBA prediction bands and MDT diagnosis. MDT, multi-disciplinary team; ASD, Autism Spectrum Disorder.

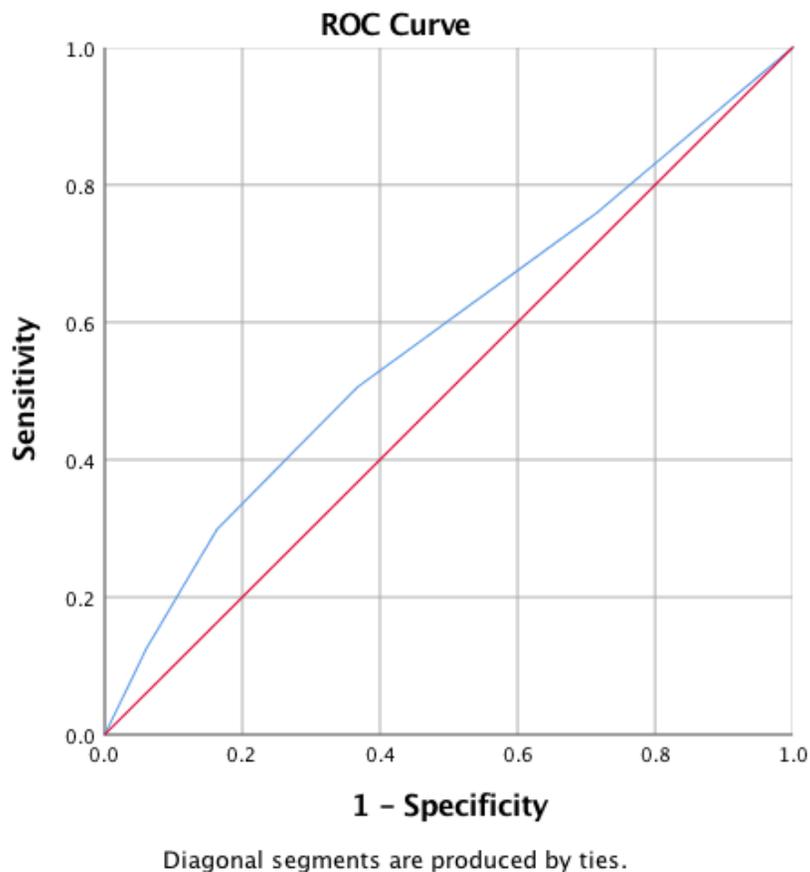


Figure 3. Receiver operative characteristic (ROC) curve of MDT diagnosis correlating with DAWBA probability bands. Area Under the Curve = 0.58 ($p = 0.13$ 95% CI = 0.5 – 0.7).

when using Consensus rating, and poor sensitivity when using computer prediction. There are several interesting findings from this study that may go towards explaining this phenomenon, to help understand the possible role for the DAWBA in future studies and/or clinical practice.

Why the unexpected result? The SCDC population

One of the main factors to consider when interpreting the results of this study is the unique nature of the children who were referred to our clinic. The mixed sensitivity and specificity results appear disappointing in this sample, but the nature of this population means that the results cannot be generalised. It is possible that there was ascertainment bias in the DAWBA interview as many of the families attending this service were seeking a second opinion (and in some cases, a third opinion). This extended contact with services may have resulted in increased mental health literacy, which may have subsequently affected their answers. This could be suspected due to the very high number of positive Consensus clinical rater diagnoses that are based on reading open-text responses and the high SDQ scores in the cohort. In addition, the ADOS (a purely observational tool) was more differentiating between ASD and non-ASD than the 3di and DAWBA which rely, in part, on parent-report.

A study by Salisbury *et al.* has highlighted that parents may already be ‘sensitised’ to ASD behaviours in children referred to a subspecialty developmental clinic compared to children in a primary care clinic. This may result in increasing the sensitivity of the interview, but not reflect its true specificity (Salisbury *et al.*, 2018). The study looked at how two screening tests performed in children between 16–48 months of age who were assessed in a subspecialty developmental clinic. The two screening tests were the Parent’s Observations of Social Interactions (POSI) and the Modified Checklist for Autism in Toddlers (M-CHAT). This clinic had a high prevalence of ASD (61%) and because of the referral process, it was suggested that parents may then be more aware of the ASD symptoms and their significance, compared to parents in the general population (Salisbury *et al.*, 2018). The high prevalence of ASD in our population with some generally very complex cases and parents’ frequent encounter with clinicians are similar. It is possible that in a more general community sample referred for first assessment of suspected autism that the findings may differ.

Another feature of our sample was that many families were seeking a second opinion in looking for answers, which suggests

a raised level of concern for autism. A study in Norway found that parents concerned about ASD are more likely to report behaviour that is consistent with ASD (Havdahl *et al.*, 2017). The authors were concerned about the diagnostic validity of the ADI-R and ADOS in Europe. Despite multiple studies and a systematic review showing high agreement of combined ADOS and ADI-R for autism diagnosis (Kim & Lord 2012, and Falkmer *et al.*, 2013, cited in Havdahl *et al.*, 2017), they were concerned the validity for both assessments were based on samples from subspecialty developmental clinics in the USA. Some studies had shown the ADI-R did not perform as well as the ADOS in a European population (Zander *et al.*, 2015, and de Bildt *et al.*, 2015, cited in Havdahl *et al.*, 2017), and the authors wondered whether factors such differences in culture and parental awareness and reporting behaviour may be impacting this (Havdahl *et al.*, 2017). Therefore, they investigated a subsample of Norwegian children that were part of a larger birth cohort study who were referred to a developmental assessment clinic. They looked at how the ADI-R and ADOS performed, with a secondary measure to see whether parental concern and reporting-behaviour contributed to the assessment outcomes (Havdahl *et al.*, 2017). This was done by dividing the sample into two groups: those children referred for possible delay, and those referred because parents were specifically concerned about possible ASD symptoms. The study demonstrated that there were higher scores on the ADI-R parent-report assessment when parents were worried about possible ASD, even independently of having a subsequent diagnosis of ASD. However, this did not affect the ADOS scores. While parents who were concerned about possible ASD were able to give examples of concerning features and symptoms (Havdahl *et al.*, 2017), this raises the question about whether they answer parent-report questionnaires in a specific way. This could be either subconsciously or consciously, but may impact on the instrument's lower validity. By the time of referral to the SCDC, the parents of our sample population have already had one MDT assessment after referral from concerned primary care health professionals who may have already raised ASD concerns. Given the above findings, it is possible that this may be implicitly affecting the way in which the DAWBA interview was completed by parents in this cohort.

The high level of parental education in our cohort may also impact the health literacy of the SCDC population. There was a very high level of high school completion in both maternal and paternal education (>70% of parents in the ASD group, >80% of parents in the non-ASD group). Studies have shown that higher socioeconomic status increases the likelihood of ASD diagnosis (King & Bearman, 2011) and that highly educated parents increases the likelihood of ASD being diagnosed earlier (Tek & Landa, 2012; Windham *et al.*, 2011). This could be associated with better access to resources, but also more health literacy or increased recognition of symptoms. This tendency to recognise symptoms and access services may have influenced the way the DAWBAs were answered and interpreted in our cohort.

Referral pathway

The majority of referrals in this cohort came from paediatricians (64% in those with ASD and 67% in those without). Again, this highlights that paediatricians found these particular cases to be complex. Specifically, these children were seen by experienced paediatricians and MDTs who were uncertain and seeking clarification of diagnosis.

The low inter-rater reliability rate

The poor inter-rater reliability from this study compared to the previously excellent inter-rater reliability calculated, is quite a stark contrast. The two unique aspects of the DAWBA in this cohort are the clinical complexity of the sample as previously discussed, and the reduction in background information provided by this version of the DAWBA. The previously reported high inter-rater reliability was based on a research project called the IMAGINE ID (Intellectual Disability and Mental Health: Assessing the Genomic Impact on Neurodevelopment). It is a research study creating a database of a children with intellectual disability, and their genotypic and phenotypic information (IMAGINE ID Study, 2018). When rating the information provided in the DAWBA by the IMAGINE ID cohort, the background risk provided significant additional information to the clinical raters, which helped contextualise the DAWBA results. This information would also have been available to the clinicians completing the MDT assessment in the clinic but was not completed by our DAWBA participants. In their initial validation study, Goodman *et al.* highlighted that clinically rated assessments enhance validity. Open-text answers allow responders to clarify their answers, and any misunderstandings around the questions or any inappropriate assumptions on typical or normative behaviour can be recognised later through the lens of a clinical rater (Goodman *et al.*, 2000). This is supported by Breslau who has argued that the whole picture of a child's situation needs to be contextualised in order to better make a diagnosis (Breslau, 1987). Future studies evaluating the DAWBA against clinical diagnosis should ensure that the background information is available to provide context and ensure validity of rating. This is particularly important in complex clinical populations.

The Strengths and Difficulties Questionnaire

The results of the Strengths and Difficulties Questionnaire show high scores and overlapping symptoms of emotional and behavioural problems between the two groups of children. The total difficulties score for both the ASD and non-ASD groups were three times higher than for the general population in the same age bracket, and children in the SCDC cohort were also nearly half as likely to have social skills. In addition, there was a significant level of conduct scores between the two groups, with levels at least three times higher than in the general population. These levels of difference raise questions about the profile of the children we assessed. It is possible that the challenge of conduct problems may be the reason for referral to this specialised clinic. It is also possible that increased conduct problems could be associated with undiagnosed psychiatric

disorder and difficult-to-manage behaviour. In addition, frustrations associated with communication problems and being misunderstood in the context of a challenging ASD diagnosis might also influence the type of population that completed the DAWBA. This is a patient group where other MDTs have had difficulty discriminating whether there was a positive ASD diagnosis, so in retrospect, expecting the DAWBA to successfully discriminate for ASD may have been an unreasonable expectation.

Comparison with ADI-Algorithm and ADOS

The results of this study in comparison to previous validation studies of the 3di and ADOS, highlight how the performance of both these assessments reflect the complexity of the SCDC population. The ADI-Algorithm and ADOS are well-validated and have been used consistently for their reliable performance. The validity study for the 3di found high sensitivity (100%) and high specificity (98%) (Skuse *et al.*, 2004) while the shortened version showed a high sensitivity (90–96%) and high specificity (85–96%) (Santosh *et al.*, 2009). Similarly, the ADOS has been demonstrated to have high validity with high sensitivity (94–100%) and moderate specificity (67–94%) (Lord *et al.*, 1994b). Charman and Gotham have discussed this in their review that results of specific screening tests are very much dependent on the sample. Factors affecting the sample include autism prevalence, the distinct features of the population, such as their age and intellect, socioeconomic and family characteristics, such as parental education, and if the assessment occurs before or after a diagnostic encounter (Charman & Gotham, 2013). They also highlight that the characteristics of the sample and the reason for screening are the two issues that affect the performance of screening assessments and the scores used to identify positive cases (Charman & Gotham, 2013). This is discussed further in the limitations section below.

The results from this study confirm that in a complex clinical group, even previously well-validated diagnostic tools can be inaccurate unless used as part of a broader MDT assessment. This further reinforces the need for rigorous assessments despite the estimated costs, and pressures faced by service providers receiving increased referrals (Galliver *et al.*, 2017).

DAWBA prediction bands

The DAWBA prediction bands function as a guide for clinicians about the risk of conditions for each participant and are helpful if clinical raters have not assessed the DAWBA. In the study used to validate the DAWBA bands, researchers found that while there were similar numbers of suggested diagnoses between the DAWBA band algorithm group and of the clinical raters, there was disagreement between the raters and computer-generated diagnoses, which were not always closely associated (Goodman *et al.*, 2010). The current study showed that no computer cut-off level had a good balance of sensitivity and specificity. The highest cut-off level had good specificity but poor sensitivity, but no level had particularly good sensitivity. The ROC curve has been used for diagnostic tests, with the area under the curve representing how accurate the test is in differentiating between those with and without the investigated

disorder (Hajian-Tilaki, 2013). Figure 3 showed that the DAWBA probability bands were not accurate in predicting ASD when compared with the gold standard in this population. Therefore, the results represented in the ROC encapsulate the overall performance of the DAWBA that was confounded by the baseline high levels of ASD and atypical presentations of the sample population. It is possible however, that in other settings, using DAWBA prediction bands may have specific utility. For example, if wanting to screen for ASD, it is possible that the lowest cut-off may provide a sensitive enough screening tool in other samples. Similarly, if there is uncertainty about a diagnosis, meeting the highest cut-off level may be considered sufficiently specific to aid in confirming a diagnosis.

The DAWBA as a possible screening measure for ASD clinics

It has been estimated that developmental assessments should have a sensitivity and specificity between 70–80% in order to be a good screening tool (American Academy of Pediatrics (Committee on Children with Disabilities), 2001). Charman and Gotham highlight that a child who is not picked up on screening in the first assessment will be picked up on any further screening if it is done intermittently (Charman & Gotham, 2013). Evidence for this is highlighted by the previously mentioned study of UK children referred to a developmental clinic, with assessments showing high sensitivity and low specificity for POSI and M-CHAT (Salisbury *et al.*, 2018). With a similar prevalence rate as our study sample (61%), the authors concluded that the low specificity was due to the clinic population, and that the tools need to be looked at in more primary care clinics (Salisbury *et al.*, 2018). However, they also highlight that clinics need to prioritise whether to use high specificity or high sensitivity to capture those at risk, which is also influenced by the amount of follow-up available (Salisbury *et al.*, 2018). Service providers may need to consider this if using the DAWBA as a screening instrument and it would depend on the amount of resources that service providers have to provide follow up assessments.

In the initial validation study for the DAWBA, Goodman *et al.* predicted that the clinical rater diagnoses would be the most useful way to implement the DAWBA as a screening test to help services plan (Goodman *et al.*, 2000). Already the DAWBA has been used in a study to confirm clinical diagnoses of ASD (Bedford *et al.*, 2017) and has been shown to have a high sensitivity and specificity of diagnosing ASD in a twin study population (McEwan & Brinkmann, 2000). Its high sensitivity in what has been proven to be a highly complex and difficult-to-diagnose clinical group, suggests that its performance in a general referral clinic could be feasible.

Charman and Gotham completed a review of screening assessments for prospective ASD diagnosis and found that these assessments are least helpful in tricky populations, where health providers need most assistance because the screening tests give less accurate results in these cases (Charman & Gotham, 2013). They do recommend the use of screening instruments in clinics, but with specific attention to the sample characteristics.

This supports the findings in this cohort that appears to contain many difficult cases, and therefore less likely to have accurate results in screening tests.

While the DAWBA as a screening test in a complex clinical population may not be helpful, it may have a role as a diagnostic adjunct. As in previous validity studies showing that the DAWBA increased the diagnosis of comorbid disorders (Goodman *et al.*, 2000; Aebi *et al.*, 2012), it is possible that the DAWBA assisted the SCDC MDT in diagnosing comorbidities. Or in differentiating ASD diagnosis from alternative diagnoses such as ADHD and anxiety. This information was not available within the scope of the study and is unfortunately a limitation. However, the data collected as part of the DAWBA, could be useful for clinicians in making decisions about children referred for further assessment in future.

Limitations

Limitations of this study include a small sample size, retrospective nature of the study, lack of follow-up, sample participant characteristics, possible clinical rater bias and limited research in the area. A larger sample size in this study may reveal more subtle differences between the DAWBA results of the ASD and non-ASD groups, or more significant evidence of difference between the two groups. A larger sample size would also allow for subgroup analysis such as looking at whether levels of parental education or referral pathway contributed to any differences.

The unique characteristics of this sample mean that it is not a diverse population. There was a high prevalence of ASD in the sample. Also, a majority had high-functioning ASD and many were difficult to diagnose and requiring a second opinion. With a more diverse sample, the results could be more generalisable. The retrospective nature of the study is also another limitation as it may have introduced selection bias, and meant a lack of follow-up. Neurodevelopmental disorders evolve with time and diagnoses can become clearer as length

of follow up time is increased. Follow-up of cases may have made diagnoses more accurate or certain and confirmed the presence or absence of comorbid disorders.

Lack of research in this area may also have impacted on the findings. The DAWBA has been validated across 17 psychopathologies but research has only been targeted at distinguishing between internalised and externalised disorders. There is a lack of focus on individual psychiatric diagnoses, with only one study to date considering ASD specifically (McEwen *et al.*, 2016). Finally, there may also be an element of bias in clinical raters. Other clinical raters might have different experiences and individual clinical biases that might make it harder to generalise the results of this study.

Conclusions

The DAWBA Consensus rating had excellent sensitivity and poor specificity in diagnosing ASD in a clinically complex referral population. These results, in such a unique population, may not be generalised to typical referral clinics. Therefore, the DAWBA may still have a role as a potential screening tool and diagnostic adjunct. Further study is required in a more general ASD referral community clinic.

Data availability

This validation study was conducted using data from patients who presented to the Social Communication Disorder Clinic at Great Ormond Street Hospital. While parents of patients gave permission for their data to be used for research, they did not consent to share this information publicly. This is in line with the recommendations by the Institutional Review Board and principle 9 of the Declaration of Helsinki. Those seeking access to the data should apply jointly to Prof. David Skuse of the Autism Families Study (d.skuse@ucl.ac.uk) and the GOSH/ICH Human Research Ethics Committee (Research.Governance@gosh.nhs.uk). Following application, consideration by these bodies will be made for allowing access to a de-identified data file.

References

- Aebi M, Kuhn C, Winkler Metzke C, *et al.*: **The use of the development and well-being assessment (DAWBA) in clinical practice: a randomized trial.** *Eur Child Adolesc Psychiatry.* 2012; **21**(10): 559–67.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Alyahri A, Goodman R: **The validation of the Arabic SDQ and DAWBA.** *East Mediterr Health J.* 2006; **12**(Suppl 2): S138–S146.
[PubMed Abstract](#)
- American Academy of Pediatrics (Committee on Children with Disabilities): **Developmental surveillance and screening of infants and young children.** *Pediatrics.* 2001; **108**(1): 192–196.
[PubMed Abstract](#) | [Publisher Full Text](#)
- American Psychiatric Association: **Diagnostic and Statistical Manual of Mental Disorders - Fifth Edition.** Washington DC, American Psychiatric Association. 2013.
[Reference Source](#)
- Bedford R, Gliga T, Shephard E, *et al.*: **Neurocognitive and observational markers: prediction of autism spectrum disorder from infancy to mid-childhood.** *Molecular Autism.* 2017; **8**: 49.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Breslau N: **Inquiring about the bizarre: false positives in Diagnostic Interview Schedule for Children (DISC) ascertainment of obsessions, compulsions, and psychotic symptoms.** *J Am Acad Child Adolesc Psychiatry.* 1987; **26**(5): 639–644.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Brett D, Warnell F, Mcconachie H, *et al.*: **Factors Affecting Age at ASD Diagnosis in UK: No Evidence that Diagnosis Age has Decreased Between 2004 and 2014.** *J Autism Dev Disord.* 2016; **46**(6): 1974–1984.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Charman T, Gotham K: **Measurement issues: Screening and diagnostic instruments for autism spectrum disorders - lessons from research and practice.** *Child Adolesc Ment Health.* 2013; **18**(1): 52–63.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- de Bildt A, Sytema S, Zander E, *et al.*: **Autism Diagnostic Interview-Revised (ADI-R) algorithms for toddlers and young preschoolers: Application in a non-US sample of 1,104 children.** *J Autism Dev Disord.* 2015; **45**(7): 2076–2091.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Department for Education (UK): **Special educational needs in England: January 2017.** Additional Tables: SFR37/2017. 2017.
[Reference Source](#)
- Dreaper J: **Autism diagnoses 'could be reduced under NHS plan'.** 2017.
[Reference Source](#)
- Falkmer T, Anderson K, Falkmer M, *et al.*: **Diagnostic procedures in autism**

spectrum disorders: a systematic literature review. *Eur Child Adolesc Psychiatry*. 2013; 22(6): 329–340.

[PubMed Abstract](#) | [Publisher Full Text](#)

Fleittlich-Bilyk B, Goodman R: **Prevalence of child and adolescent psychiatric disorders in southeast Brazil.** *J Am Acad Child Adolesc Psychiatry*. 2004; 43(6): 727–34.

[PubMed Abstract](#) | [Publisher Full Text](#)

Ford T, Goodman A, Meltzer H: **The British Child and Adolescent Mental Health Survey 1999: the prevalence of DSM IV disorders.** *J Am Acad Child Adolesc Psychiatry*. 2003; 42(10): 1203–1211.

[PubMed Abstract](#) | [Publisher Full Text](#)

Galliver M, Gowling E, Farr W, *et al.*: **Cost of assessing a child for possible autism spectrum disorder? An observational study of current practice in child development centres in the UK.** *BMJ Paediatrics Open*. 2017; 1(1): e000052.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Gilbert R, Logan S, Moyer V, *et al.*: **Assessing diagnostic and screening tests: Part 1. Concepts.** *West J Med*. 2001; 174(6): 405–9.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Goodman A, Heiervang E, Collishaw S, *et al.*: **The 'DAWBA bands' as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples.** *Soc Psychiatry Psychiatric Epidemiology*. 2010; 46(6): 521–532.

[PubMed Abstract](#) | [Publisher Full Text](#)

Goodman R: **Scoring instructions for SDQs for 4-17 year olds, completed by parents, teachers or self-report.** United Kingdom: youthmind. 2018.

[Reference Source](#)

Goodman R, Ford T, Richards H, *et al.*: **The development and well-being assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology.** *J Child Psychol Psychiatry*. 2000; 41(5): 645–655.

[PubMed Abstract](#)

Great Ormond Street Hospital for Children NHS: Research [Online]. London, U.K.: Great Ormond Street Hospital for Children. 2020.

[Reference Source](#)

Hajian-Tilaki K: **Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation.** *Caspian J Intern Med*. 2013; 4(2): 627–635.

[PubMed Abstract](#) | [Free Full Text](#)

Havdahl KA, Bishop SL, Surén P, *et al.*: **The influence of parental concern on the utility of autism diagnostic instruments.** *Autism Res*. 2017; 10(10): 1672–1686.

[PubMed Abstract](#) | [Publisher Full Text](#)

Heiervang E, Stormark KM, Lundervold AJ, *et al.*: **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*. 2007; 46(4): 438–447.

[PubMed Abstract](#) | [Publisher Full Text](#)

IBM Corp: **IBM SPSS Statistics for Windows.** 25.0 ed. Armonk, NY: IBM Corp. 2017.

[Reference Source](#)

Imagine id study: **Intellectual Disability and Mental Health: Assessing Genomic Impact on Neurodevelopment - Report on Progress: February 2017 to February 2018.** London, United Kingdom: Institute of Child Health: University College London. 2018

[Reference Source](#)

Kim SH, Lord C: **Combining information from multiple sources for the diagnosis of autism spectrum disorders for toddlers and preschoolers from 12 to 47 months of age.** *J Child Psychol Psychiatry*. 2012; 53(2): 143–151.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

King M, Bearman P: **Socioeconomic Status and the Increased Prevalence of Autism in California.** *Am Sociol Rev*. 2011; 76(2): 320–346.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Lord C, Risi S, Lambrecht L, *et al.*: **The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associate with the spectrum of autism.** *Journal of Autism and Developmental Disorders*. 2000; 30(3): 205–223.

[PubMed Abstract](#) | [Publisher Full Text](#)

Lord C, Rutter M, Le Couteur A: **Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders.** *J Autism Dev Disord*. 1994a; 24(5): 659–685.

[PubMed Abstract](#) | [Publisher Full Text](#)

Lord C, Rutter M, Le Couteur A: **The Autism Diagnostic Interview-Revised: A Revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders.** *J Autism Dev Disord*. 1994b; 24(5): 659–685.

[PubMed Abstract](#) | [Publisher Full Text](#)

Mannon A, Leader G: **Comorbidity in autism spectrum disorder: A literature review.** *Research in Autism Spectrum Disorders*. 2013; 7(12): 1595–1616.

[PubMed Abstract](#)

Mcewen JJ, Brinkmann AO: **Androgen Physiology: Receptor and Metabolic Disorders.** De Groot LJ, Chrousos G, Dungan K, Grossman A, Hershman JM, Koch C, Korbonits M, McLachlan R, New M, Purnell J, Rebar R, Singer F, Vinik A. (eds.) *Endotext*. South Dartmouth MA: MDText.com, Inc. 2000.

[PubMed Abstract](#)

Mcewen FS, Stewart CS, Colvert E, *et al.*: **Diagnosing autism spectrum disorder in community settings using the Development and Well-Being Assessment: validation in a UK population-based twin sample.** *J Child Psychol Psychiatry*. 2016; 57(2): 161–170.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Meltzer H, Gatward R, Goodman R, *et al.*: **Mental health of children and adolescents in Great Britain.** London: The Stationery Office: Social Survey Division of the Office for National Statistics on behalf of the Department of Health, the Scottish Health Executive and the National Assembly for Wales. 2000.

[PubMed Abstract](#)

Meltzer H, Gatward R, Goodman R, *et al.*: **Mental health of children and adolescents in Great Britain.** *Int Rev Psychiatry*. 2003; 15(1–2): 185–187.

[PubMed Abstract](#) | [Publisher Full Text](#)

Mullick M, Goodman R: **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*. 2005; 40(8): 663–71.

[PubMed Abstract](#) | [Publisher Full Text](#)

NICE: **Autism Diagnosis in children and young people: Recognition, referral and diagnosis of children and young people on the autism spectrum.** 2011. Clinical Guidance 128. London: National Institute for Health and Care Excellence.

[PubMed Abstract](#)

Oosterling I, Roos S, de Bildt A, *et al.*: **Improved Diagnostic Validity of the ADOS Revised Algorithms: A Replication Study in an Independent Sample.** *J Autism Dev Disord*. 2010; 40(6): 689–703.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Salisbury L, Nyce J, Charles D, *et al.*: **Sensitivity and Specificity of 2 Autism Screeners Among Referred Children Between 16 and 48 Months of Age.** *J Dev Behav Pediatr*. 2018; 39(3): 254–258.

[PubMed Abstract](#) | [Publisher Full Text](#)

Santosh P, Mandy W, Puura K, *et al.*: **The construction and validation of a short form of the developmental, diagnostic and dimensional interview.** *European Child and Adolescent Psychiatry*. 2009; 18(8): 521–524.

[PubMed Abstract](#) | [Publisher Full Text](#)

Simonoff E, Pickles A, Charman T, *et al.*: **Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample.** *J Am Acad Child Adolesc Psychiatry*. 2008; 47(8): 921–929.

[PubMed Abstract](#) | [Publisher Full Text](#)

Skuse D, Warrington R, Bishop D, *et al.*: **The developmental, dimensional and diagnostic interview (3di): A novel computerized assessment for autism spectrum disorders.** *J Am Acad Child Adolesc Psychiatry*. 2004; 43(5): 548–558.

[PubMed Abstract](#) | [Publisher Full Text](#)

Soke G, Maenner M, Christensen D, *et al.*: **Prevalence of Co-occurring Medical and Behavioral Conditions/Symptoms Among 4- and 8-Year-Old Children with Autism Spectrum Disorder in Selected Areas of the United States in 2010.** *J Autism Dev Disord*. 2018; 48(8): 2663–2676.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Tek S, Landa R: **Differences in Autism Symptoms between Minority and Non-Minority Toddlers.** *J Autism Dev Disord*. 2012; 42(9): 1967–1973.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

The UK National Screening Committee: **Chapter Two: Definitions and Classifications of Population Screening Programmes. First Report of the National Screening Committee.** London: Department of Health. 1998.

[Reference Source](#)

Wilson J, Jungner G: **Principles and practice of screening for disease.** 1968.

[Reference Source](#)

Windham G, Anderson M, Croen L, *et al.*: **Birth Prevalence of Autism Spectrum Disorders in the San Francisco Bay Area by Demographic and Ascertainment Source Characteristics.** *J Autism Dev Disord*. 2011; 41(10): 1362–72.

[PubMed Abstract](#) | [Publisher Full Text](#)

Zander E, Sturm H, Bolte S: **The added value of the combined use of the Autism Diagnostic Interview-Revised and the Autism Diagnostic Observation Schedule: Diagnostic validity in a clinical Swedish sample of toddlers and young preschoolers.** *Autism*. 2015; 19(2): 187–199.

[PubMed Abstract](#) | [Publisher Full Text](#)

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research