Published in final edited form as:

J Autism Dev Disord. 2015 March ; 45(3): 873–880. doi:10.1007/s10803-014-2187-6.

Brief Report: Independent Validation of Autism Spectrum Disorder Case Status in the Utah Autism and Developmental Disabilities Monitoring (ADDM) Network Site

Amanda V. Bakian,
Department of Psychiatry, University of Utah, 650 Komas Drive Suite 206, Salt Lake City, UT 84108, USA

Deborah A. Bilder,
Department of Psychiatry, University of Utah, 650 Komas Drive Suite 206, Salt Lake City, UT 84108, USA

Paul S. Carbone,
Department of Pediatrics, University of Utah, Salt Lake City, UT, USA

Tyler D. Hunt,
Department of Psychiatry, University of Utah, 650 Komas Drive Suite 206, Salt Lake City, UT 84108, USA

Department of Pediatrics, University of Utah, Salt Lake City, UT, USA

Brent Petersen, and
Valley Mental Health's Carmen B. Pingree Center for Children with Autism, Salt Lake City, UT, USA

Catherine E. Rice
National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, GA, USA

Amanda V. Bakian: amanda.bakian@hsc.utah.edu

Abstract

An independent validation was conducted of the Utah Autism and Developmental Disabilities Monitoring Network's (UT-ADDM) classification of children with autism spectrum disorder (ASD). UT-ADDM final case status (n = 90) was compared with final case status as determined by independent external expert reviewers (EERs). Inter-rater reliability (ICC = 0.84), specificity [0.83 (95 % CI 0.74–0.90)], and sensitivity [0.99 (95 % CI 0.96–1.00)] were high for ASD case versus non-case classification between UT-ADDM and EER. At least one EER disagreed with UT-ADDM on ASD final case status on nine out of 30 records; however, all three EERs disagreed with UT-ADDM for only one record. Findings based on limited data suggest that children with

Correspondence to: Amanda V. Bakian, amanda.bakian@hsc.utah.edu.

This study is dedicated to our friend and co-author, Dr. Brent Petersen who passed away prior to the completion of this manuscript. Brent dedicated his life to serving others including children with autism and their families.
Brent Petersen: Deceased.

ASD as identified by UT-ADDM are consistently classified as ASD cases by independent autism experts.

Keywords

Autism spectrum disorder; Autism and Developmental Disabilities Monitoring Network; ASD surveillance; ASD prevalence; Validation; DSM-IV-TR

Introduction

The identified prevalence of autism spectrum disorders (ASD) has increased significantly over the last decade according to the Centers for Disease Control and Prevention's Autism and Developmental Disabilities Monitoring (ADDM) Network (Centers for Disease Control and Prevention 2007a, b, 2009, 2012, 2014). The ADDM Network's most recent estimate identified 1.47 % of 8-yearold children (one in 68) with an ASD in 2010 (Centers for Disease Control and Prevention 2014). A strength of the ADDM Network surveillance approach is that prevalence estimates are derived from a common methodology which is applied across study years by participating sites (Van Naarden Braun et al. 2007), thereby minimizing comparison errors based on variations in methodology. The approach uses a multisource, retrospective records review of existing health records (e.g. diagnostic and developmental assessments from psychologists, child psychiatrists, neurologists, developmental pediatricians, physical therapists, occupational therapists, speech/language pathologists) and, when available, education records. The Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, text revision (*DSM-IV-TR*) (American Psychiatric Association 2000) is used to define ASD case status. Children with ASD are identified according to *DSM-IV-TR* criteria through an independent clinician review of the developmental information in the records. An evaluation of the approach suggests that it is a cost-effective alternative to conducting a complete population screening and diagnostic clinical assessment of ASD (Van Naarden Braun et al. 2007). In a validation study, Avchen et al. (2011) reported fairly high specificity [0.96 (95 % CI 0.94–0.99)], positive predictive value [0.79 (95 % CI 0.66–0.93)], and negative predictive value [0.91 (95 % CI 0.87–0.96)] of the ADDM surveillance methodology compared to direct clinical assessment. The low sensitivity measured in the study [0.60 (95 % CI 0.45–0.75)] suggests that prevalence estimates derived from the ADDM-based ASD surveillance approach are conservative.

Despite the effort to maintain consistent methodology over time within and across ADDM sites, significant variation has been measured in ASD prevalence estimates between sites. In 2008, the Utah ADDM site (UT-ADDM) reported the highest estimated ASD prevalence (21.2 per 1,000 or 1 in 47) among all 14 participating sites (Centers for Disease Control and Prevention 2012), significantly exceeding prevalence estimates at all but two other sites (Arizona and New Jersey).

The complex nature of ASD, the absence of a biological marker, changes in diagnostic practices over time, and variation in the availability of diagnostic services complicates the pursuit of finding accurate ASD prevalence rates (Liptak et al. 2008). However, it is paramount that community diagnosticians, stakeholders and the general public have trust in

the accuracy of the prevalence estimated by public health entities. In light of these challenges and UT-ADDM's comparatively-high estimated ASD prevalence, a small validation study was conducted to compare ASD final case status as determined by UT-ADDM with ASD case status as determined by community provider, independent external expert reviewers (EERs). This study's objectives were: (1) to measure agreement between UT-ADDM and EERs on final case status, ASD subtype classification (autism vs. ASD-NOS), degree of case certainty, and degree of case impairment; and (2) to examine characteristics of the records on which there was disagreement in ASD final case status between UT-ADDM and EERs.

Methods

ADDM Network Surveillance Approach

In study year 2008 (SY2008), the UT-ADDM site implemented a surveillance methodology common to the ADDM network that is based on the CDC's Metropolitan Atlanta Developmental Disabilities Surveillance Program (MADDSP) model to identify children aged 8 years with ASD. The MADDSP and ADDM surveillance methods have been described in detail elsewhere (Centers for Disease Control and Prevention 2007a, b, 2009; Rice et al. 2007; Van Naarden Braun et al. 2007; Yeargin-Allsopp et al. 2003). ASD case determination was conducted in two phases: (1) screening and abstraction of records at multiple data sources in the community; and (2) de-identified abstracted composite evaluation files were reviewed by trained clinicians to determine ASD final case status.

Phase I: Screening and Abstraction

In the first phase, evaluation records from a population of 8-year-old children residing in a continuous geographical region of Utah (n = 2,123) were screened to identify a subset of children previously identified with ASD, suspected of having ASD, or exhibiting social symptoms of ASD. Multiple types of data records were screened including (1) educational records that described eligibility for special education services, and (2) health records that might have included diagnostic and developmental assessments from a variety of providers (e.g., psychologists, child psychiatrists, developmental pediatricians, educational specialists). Children's records were screened to confirm that the children belonged to the 2000 birth cohort and were residents of the surveillance area at age eight in 2008. Next, eligible records were screened for ASD "triggers" which included any mention of a community diagnosis of autism or social behaviors associated with ASD (e.g., prefers to play alone, does not make eye contact). Following the identification of an ASD trigger, developmental information from birth through age eight was abstracted (n = 613 health evaluations, n = 458 educational evaluations) and merged into a single composite record for each child (n = 151). Ten evaluations were abstracted on average per child.

Phase II: Clinician Review

In the second phase, the composite record was reviewed by two UT-ADDM CDC-trained clinician reviewers to determine ASD final case status. ASD final case status was operationalized into three categories: Confirmed Case, Suspected Case, or Does Not Qualify (DNQ). A child met the ADDM case definition of Confirmed Case if he or she displayed

behaviors documented in the composite evaluation that were consistent with the *DSM-IV-TR* criteria for Autistic Disorder, Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS) including Atypical Autism, or Asperger's Disorder. For children meeting the surveillance case definition of Confirmed Case, two subtype classifications were operationalized: Autistic Disorder or Autism Spectrum Disorder-Not Otherwise Specified (ASD-NOS). Children meeting the definition of the Autistic Disorder subtype exhibited at least six *DSM-IV-TR* criteria distributed across at least two social, one communication, and one behavior criteria. Evidence was also required of a developmental delay preceding 3 years of age. Children meeting the definition of the ASD-NOS subtype exhibited at least one social criterion and one communication or behavior criterion. For ASD-NOS, an additional feature considered indicative of ASD was required beyond the *DSM-IV-TR* criteria, such as loss of social or language skills, seeming oblivious to adults or others in the presence of a clear social opportunity, or experiencing a consistent preoccupation with a narrow and unusual interest.

In addition to coding the number and pattern of *DSM-IV-TR* criteria present in the composite records, UT-ADDM clinician reviewers rated the record's overall quality (scale of 1–5 from poor to excellent). Record quality was judged on the amount and detail of information in the record that could be used to confirm or rule out ASD. For Confirmed or Suspected cases, UT-ADDM clinician reviewers also recorded their impression of the child's level of impairment (scale of 1–5 from mild to severe), and their degree of certainty (scale of 1–5 from not sure to very sure). A degree of certainty score of at least "4" was required from the primary reviewer for a child to be included as an ASD case without further review. If a child met the ADDM ASD case definition but the clinician reviewer's degree of certainty was between "1" ("not sure") and "3" ("somewhat sure"), then a second review was conducted by a UT-ADDM clinician reviewer. Following the completion of both reviews, the first and second reviewers discussed the record and established consensus on the record's final case status and additional measures including degree of certainty, level of impairment, and overall record quality.

Selection of Records and External Expert Review

Thirty records with a final case status of ASD (subtype classification: $n = 26$ Autistic Disorder: $n = 4$ ASD-NOS), 30 records with a final case status of suspected ASD, and 30 records with a final case status of DNQ were randomly selected for EERs from the 151 total composite records reviewed by UT-ADDM clinicians in SY2008. The EERs did not directly assess in person the children from whom these records were derived. External expert reviewers included a child psychiatrist (EER1), a child psychiatrist/pediatrician (EER2), and a pediatrician (EER3), who all had long-term clinical experience in the field of autism and are considered experts in diagnosing ASD. The EERs, blinded to ADDM final case status, evaluated the 90 randomly selected UT-ADDM composite records and reported their clinical impressions on a coding form developed for this study (see "Appendix"). They were requested to determine final case status (Confirmed Case, Suspected Case, or DNQ), and ASD subtype (Autistic Disorder or ASD-NOS). EERs also recorded quality of the composite record using a rating scale similar to the ADDM scale (1–5 from poor to excellent). For final case status of suspected or DNQ, EERs were asked to specify the most applicable reason for

their decision from the following: insufficient information, conflicting information, could be or clearly could be accounted for by another disorder, sufficient information to rule out an ASD, or other reason. If the reported behaviors could be explained by an alternative disorder(s), the EERs were asked to specify. No secondary or consensus review process was implemented among EERs.

Statistical Analysis

UT-ADDM versus EER final case status inter-rater reliability was measured using intraclass correlations (ICC). Agreement was reached on final case status between UT-ADDM and EER when the final case status of all three EERs matched UT-ADDM's final case status. When the final case status of at least one EER differed from UT-ADDM's final case status, the case was considered to be discordant. ASD final case status was also collapsed into two categories: case (Confirmed case) versus non-case (Suspected and DNQ). Inter-rater reliability was measured for three-level final case status (Confirmed case vs. Suspected case vs. DNQ), two-level final case status (case vs. non-case), non-case final case status (Suspected case vs. DNQ), and ASD subtype (Autistic Disorder vs. ASDNOS). The clinical sensitivity, specificity, positive predictive value (PPV), and negative predicative value (NPV) of EER review using UT-ADDM as the gold standard were estimated for two-level final case status (case vs. non-case) from a 2 × 2 table. These values are referred to as clinical estimates of sensitivity, specificity, and predicted value because the comparison is between two clinician review approaches and not, as is traditional, two tests. Differences in characteristics between Confirmed cases with discordant agreement between EER and UT-ADDM were compared using Fisher's Exact test for categorical variables and *t* tests for continuous variables. Analyses were conducted in the R system for statistical computing (R Core Team 2011) or in SAS version 9.3 (SAS Institute, Cary NC). An alpha level of 0.05 was selected to indicate significance for all statistical tests.

Results

Table 1 reports the inter-rater reliability calculations between the external expert reviewers and UT-ADDM. Inter-rater reliability was high between UT-ADDM and EER for two-level final case status (case vs. non-case): F (89, 270) = 21.6, ICC = 0.84 (95 % CI 0.79–0.88), $p < 0.001$. Whereas, poor agreement was measured for three-level final case status (Confirmed case vs. Suspected case vs. DNQ): F (89, 270) = 6.19, ICC = 0.57 (95 % CI 0.47–0.66), $p < 0.001$. Poor agreement was measured for ASD subtype classification (Autistic Disorder vs. ASDNOS): F (20, 63) = 2.64, ICC = 0.29 (95 % CI 0.08–0.55), $p = 0.002$, and agreement was null for noncase classification (Suspected Case vs. DNQ) between UT-ADDM and EER: F (57, 174) = 0.89, ICC = −0.03 (95 % CI −0.11–0.09), $p = 0.69$. Overall, measures of clinical sensitivity [0.70 (95 % CI 0.51–0.85)], specificity [0.97 (95 % CI 0.88–1.00)], positive predicted value [0.91 (95 % CI 0.72–0.99)], and negative predicted value [0.87 (95 % CI 0.76–94)] were fairly high.

At least one of the EERs disagreed with UT-ADDM on ASD final case status in nine out of 30 records that UT-ADDM classified as a Confirmed case. This represents a fair bit of disagreement across individual EERs. However, for only one Confirmed case did all three of

the EERs disagree with UT-ADDM. For this child, all EERs classified the child's status as "Suspected ASD" with multiple other disorders (i.e., intellectual disability, mood disorder, language delays) creating a complicated diagnostic picture. Table 2 presents composite record-level characteristics of the 21 records with a UT-ADDM final case status of Confirmed case for which there was perfect agreement between UT-ADDM and all three EERs, and the nine records with a UT-ADDM final case status of Confirmed case for which there was disagreement between UT-ADDM and at least one EER. Twenty-five out of the 30 children that were Confirmed ADDM ASD cases were male. Nineteen out of 21 (90 %) of the Confirmed cases for which there was agreement between UT-ADDM were male and six out of nine (67 %) of the Confirmed cases for which there was disagreement were male. Of the 21 Confirmed cases for which there was agreement between UT-ADDM and EERs, 62 % had a previous diagnosis of ASD on record and 43 % had an autism special education eligibility on record. Only 22 % of the nine Confirmed cases for which there was disagreement between UT-ADDM and EER had a previous ASD diagnosis on record and only 11 % had an autism special education eligibility. However, no statistically significant differences in characteristics were found between Confirmed cases for which there was agreement versus disagreement between EER and UT-ADDM (Table 2). In addition, there were no differences in mean IQ score between the Confirmed cases on which there was UT-ADDM and EER agreement (mean IQ = 75.4, SD = 17.4) versus the Confirmed cases on which there was UT-ADDM and EER disagreement (mean IQ = 64.5, SD = 22.14; p = 0.23). "Other disorder(s)" was indicated by at least one of the EERs on eight out of nine discordant records as a reason for not confirming the child as an ASD case. Alternative health conditions suggested by the EERs to explain the child's observed behavior included intellectual/cognitive disability (5/9), anxiety (3/9), or language impairment (2/9). Insufficient information was also indicated as a reason for not confirming the child as an ASD case on eight out of nine discordant records. The average record quality was 4.16 for the 21 UT-ADDM Confirmed cases for which there was agreement between UT-ADDM and EER while the average record quality was 2.89 for the nine UT-ADDM Confirmed cases on which UT-ADDM and EER disagreed. Finally, EERs and UT-ADDM agreed on final case status for 58/60 non-cases (includes Suspected and DNQ). The two children confirmed as ASD cases by EER but not UT-ADDM were both classified as an ASD case by only one out of three EERs.

Discussion

Two-level final case status (case vs. non-case) inter-rater agreement and corresponding clinical specificity, sensitivity, positive predictive value, and negative predictive value were fairly high between UT-ADDM and the three EERs providing convergent validation of the ADDM method for identifying children with ASD. In our study, among UT-ADDM Confirmed cases, at least one EER agreed with UT-ADDM clinician review on ASD final case status in 29 out of 30 records. Disagreement among EERs on final case status determination in nine out of 30 ADDM identified children with ASD indicates a somewhat high level of variation across individual EERs. Reasons cited by the EERs for not confirming the nine discordant cases as ASD included insufficient information and the presence of other disorders. Discordant records were also of low average quality compared

to the 21 Confirmed cases on which there was agreement between UT-ADDM and the three EERs.

While two-level final case status inter-rater reliability between UT-ADDM and EER was excellent, characteristics of Utah's ASD cases could make it difficult for EERs to confirm ASD case status through a retrospective record review process alone. Several key characteristics of UT-ADDM ASD cases differ from ASD cases identified at other SY2008 ADDM sites (Centers for Disease Control and Prevention 2012) and from cases described in previously published ASD studies (e.g. Fombonne 2009). Utah, along with the New Jersey ADDM site, identified a higher proportion of children with intellectual quotients (IQ) above 70 compared with any of the other sites participating in the ADDM network in SY2008. Eighty-seven percent of the UT-ADDM cases had an IQ >70 compared with 62 % of cases across all ADDM sites combined. Although ASD prevalence estimates were significantly higher for boys than girls among all ADDM sites combined ($p < 0.01$), the male-to-female prevalence ratio was lower in Utah (2.7) compared with all ADDM sites combined (4.6). Despite some differences in the characteristics of Utah's ASD cases compared with other ADDM sites, no statistically significant differences in sex and average IQ were detected between Confirmed cases for which there was agreement and Confirmed cases for which there was disagreement between UT-ADDM and EER. This, as well as the high two-level final case status agreement measured between UT-ADDM and EER, suggests that the ADDM method is robust to variation in underlying case characteristics for ascertaining ASD.

We were not surprised to measure poor inter-rater reliability for ASD subtype (Autistic Disorder vs. ASD-NOS) between UT-ADDM and EER. The ADDM record review methodology is not intended to classify clinically-meaningful ASD subtypes, although the diagnostic subtypes indicated by community providers on comprehensive evaluations were reported in publications of prevalence findings (Centers for Disease Control and Prevention 2009, 2012, 2014). Even among research and clinical practices involving highly trained ASD diagnosticians distinguishing among ASD subtypes (Autistic Disorder, PDD-NOS, and Asperger's Disorder) is known to be inconsistent (Lord et al. 2012; Risi et al. 2006). Changes to ASD classification in the newest iteration of the American Psychiatric Association's ASD diagnostic criteria reflect this finding (American Psychiatric Association 2013).

An evaluation of differences in characteristics between records from UT-ADDM Confirmed cases for which there was agreement versus disagreement between UT-ADDM and EER found that a larger proportion of records on which there was agreement had previous indications of an ASD diagnosis and/or an autism special education classification, although these differences did not reach statistical significant. Discordant records were of lower average quality than records on which there was agreement. In addition, EERs cited insufficient information and the presence of another disorder as reasons for not classifying a child as a case. Overall, each individual EER identified fewer ASD cases than UT-ADDM. The reasons, cited above, that EERs provided for not confirming a child as an ASD case suggest that expert clinician review of retrospective records was conservative relative to UT-ADDM.

The measured prevalence of ASD varied widely across ADDM sites in 2008 (Centers for Disease Control and Prevention 2012). Although this study focuses on a single site that participated in the ADDM network in 2008, our findings concerning disagreement between EER and ADDM clinician reviewers may shed light on potential reasons for this phenomenon. Our study found that insufficient information and record quality influenced the EER's determination of final case status. The amount and quality of the information contained in a child's composite record is known to vary across ADDM sites (Van Naarden Braun et al. 2007) and has been suggested as an important factor driving variation and change in the measured prevalence of ASD in the ADDM network (Rice et al. 2010). In addition, disagreement in final case status determination among individual EERs and between EER and UT-ADDM may reflect the variability that exists, even among clinical experts practicing within the same community, in ASD diagnostic threshold based on a records review process. As such, this study's findings highlight the importance of having a clear and consistent records review process shared across ADDM sites, the fidelity of which is maintained by regular reliability exercises.

A number of researchers have suggested that increases in ASD prevalence estimates are a result, in part, of diagnostic substitution and accretion (i.e., increasingly diagnosing autism in addition to other conditions previously diagnosed) (Bishop et al. 2008; King and Bearman 2009; Shattuck 2006a, b). Although not the focus of this study, several of the discordant cases involved children for which co-morbid intellectual disability, language impairment, sensory and/or psychiatric disorders were noted by the EERs as alternative health conditions. This indicates that clinical complexity may have impacted the EER's case determination.

This study has several strengths including the evaluation of data collected in Utah as part of a population-based assessment of ASD prevalence. The strong agreement measured between UT-ADDM and EER on ASD final case status provides validation of the ADDM method to identify ASD cases which serve as the basis for estimating ASD prevalence in Utah. As ongoing clinician review quality control procedures are conducted among ADDM sites to enable comparisons of final case status across the ADDM Network, our Utah validation findings could have network-wide implications toward providing confidence in clinician review determination of ASD final case status at other ADDM sites. At the same time, some weaknesses are noted, including this study's relatively small sample size. Also, validation was conducted using retrospective review of pre-existing records instead of via an independent clinical exam as was conducted in a previous ADDM Network validation study (Avchen et al. 2011).

Obtaining accurate estimates of ASD prevalence is imperative for determining the overall public health impact of ASD and planning for future diagnostic and treatment services to assist individuals with this increasingly identified condition. The approach used in this study was a means to validate UT-ADDM ASD final case status as identified using ADDM methodology. Overall, our findings suggest that children classified as ASD cases in the ADDM Network are also likely to be classified as ASD cases by independent experts, and provides support for UT-ADDM ASD prevalence estimates derived using the ADDM methodology.

Acknowledgments

Surveillance data used for this study were collected under the Grant/Cooperative Agreement UR3DD000685 from the Centers for Disease Control and Prevention. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Appendix

ASD Independent External Expert Review Summary Coding Form

ASD INDEPENDENT REVIEW SUMMARY CODING FORM

STUDYID: _____ Review Date: _____ Time to Review: _____

Reviewer: _____

Reviewer rating of quality of record: 1 2 3 4 5

Poor Adequate Excellent

Final ASD Case Definition:

_____ **Confirmed ASD Case+:** *ASD Review Classification:* _____*Autism* _____*ASD-NOS*

+Reviewer Degree of Certainty that child is an ASD case:

1* 2* 3* 4 5

Not sure Somewhat Sure Very Sure

+Reviewer Degree of Impairment Associated with ASD:

1 2 3 4 5

Mild Moderate Severe

_____ **Suspected Case+:** _____*Probable ASD case^** _____*Possible ASD case**

+Reviewer Degree of Certainty that child is an ASD case:

1* 2* 3* 4^* 5^*

Not sure Somewhat Sure Very Sure

_____ **DNQ After Review ***

*** Specify most applicable reason if a confirmed case with Certainty (1-3); Suspected Case; or DNQ After Review**
_____ Insufficient information
_____ Conflicting Information
_____Could be accounted for by another disorder (s) (specify): _____
_____Clearly accounted f or by other disorders (s) (specify): _____
_____Sufficient Information to rule out an ASD (specify): _____
_____Other (specify): _____

References

American Psychiatric Association. Diagnostic criteria from DSM-IV-TR. Washington, DC: American Psychiatric Association; 2000.

American Psychiatric Association. Diagnostic criteria from DSM-5. Washington, DC: American Psychiatric Association; 2013.

Avchen RN, Wiggins LD, Devine O, Van Naarden Braun K, Rice C, Hobson NC, et al. Evaluation of a records-review surveillance system used to determine the prevalence of autism spectrum disorders. Journal of Autism and Developmental Disorders. 2011; 41(2):227–236. [PubMed: 20568003]

Bishop DV, Whitehouse AJ, Watt HJ, Line EA. Autism and diagnostic substitution: Evidence from a study of adults with a history of developmental language disorder. Developmental Medicine and Child Neurology. 2008; 50(5):341–345. [PubMed: 18384386]

Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders: Autism and Developmental Disabilities Monitoring Network, six sites, United States, 2000. MMWR Surveillance Summaries. 2007a; 56(1):1–11.

Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders: Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2002. MMWR Surveillance Summaries. 2007b; 56(1):12–28.

Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders: Autism and Developmental Disabilities Monitoring Network, United States, 2006. MMWR Surveillance Summaries. 2009; 58(10):1–20.

Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders: Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. MMWR Surveillance Summaries. 2012; 61(3):1–21.

Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders among children aged 8 years: Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2010. MMWR Surveillance Summaries. 2014; 63(2):1–19.

Fombonne E. Epidemiology of pervasive developmental disorders. Pediatric Research. 2009; 65(6): 591–598. [PubMed: 19218885]

King M, Bearman P. Diagnostic change and the increased prevalence of autism. International Journal of Epidemiology. 2009; 38(5):1224–1234. [PubMed: 19737791]

Liptak GS, Benzoni LB, Mruzek DW, Nolan KW, Thingvoll MA, Wade CM, et al. Disparities in diagnosis and access to health services for children with autism: Data from the National Survey of Children's Health. Journal of Developmental and Behavioral Pediatrics. 2008; 29(3):152–160. [PubMed: 18349708]

Lord C, Petkova E, Hus V, Gan W, Lu F, Martin DM, et al. A multisite study of the clinical diagnosis of different autism spectrum disorders. Archives of General Psychiatry. 2012; 69(3):306–313. [PubMed: 22065253]

R Core Team. A language and environment for statistical computing. Viena, Austria: R Foundation for Statistical Computing; 2011. ISBN 3-900051-07-0

Rice CE, Baio J, Van Naarden Braun K, Doernberg N, Meaney FJ, Kirby RS. A public health collaboration for the surveillance of autism spectrum disorders. Paediatric and Perinatal Epidemiology. 2007; 21(2):179–190. [PubMed: 17302648]

Rice CE, Nicholas J, Baio J, Pettygrove S, Lee L, Van Naarden Braun K, et al. Changes in autism spectrum disorder prevalence in 4 areas of the United States. Disability and Health Journal. 2010; 3(3):186–201. [PubMed: 21122784]

Risi S, Lord C, Gotham K, Corsello C, Chrysler C, Szatmari P, et al. Combining information from multiple sources in the diagnosis of autism spectrum disorders. Journal of the American Academy of Child and Adolescent Psychiatry. 2006; 45(9):1094–1103. [PubMed: 16926617]

Shattuck PT. The contribution of diagnostic substitution to the growing administrative prevalence of autism in US special education. Pediatrics. 2006a; 117(4):1028–1037. [PubMed: 16585296]

Shattuck PT. Diagnostic substitution and changing autism prevalence. Pediatrics. 2006b; 117(4):1438–1439. [PubMed: 16585346]

Van Naarden Braun K, Pettygrove S, Daniels J, Miller L, Nicholas J, Baio J, et al. Evaluation of a methodology for a collaborative multiple source surveillance network for autism spectrum disorders: Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2002. MMWR Surveillance Summaries. 2007; 56(1):29–40.

Yeargin-Allsopp M, Rice C, Karapurkar T, Doernberg N, Boyle C, Murphy C. Prevalence of autism in a US metropolitan area. JAMA. 2003; 289(1):49–55. [PubMed: 12503976]

Table 1

Intraclass correlations (ICC) and 95 % confidence intervals (95 % CI) between independent external expert reviewers (EERs) and Utah Autism and Developmental Disabilities Monitoring (UT-ADDM) clinician review

Test	ICC	95 % CI	*p* value
Case versus non-case[a]	0.84	0.79–0.88	<.001
Case versus Suspected case versus DNQ[b]	0.57	0.47–0.66	<.001
Suspected case versus DNQ[b]	−0.03	−0.11–0.09	0.69
Degree of certainty (ordinal scale 1–5)[c]	0.55	0.34–0.76	<.001
Degree of impairment (ordinal scale 1–5)[c]	0.08	−0.08–0.33	0.18
ASD subtype (Autistic Disorder vs. ASD-NOS)[c]	0.29	0.08–0.55	0.002

[a] Includes Suspected and DNQ

[b] *DNQ* Does Not Qualify

[c] Cases only

Table 2

Characteristics of records from 30 UT-ADDM Confirmed cases as a function of two-level UT-ADDM versus EER final case status agreement

Measure	Agreement *n* (%)[a]		*p* value[b]
	Yes (*n* = 21) (%)	No (*n* = 9) (%)	
Previous ASD diagnosis[c]	13 (62)	2 (22)	0.11
Special education[d]	15 (71)	7 (78)	0.99
Autism special education[e]	9 (43)	1 (11)	0.11
Male	19 (90)	6 (67)	0.14
Reason for EER non-case decision[f]			
Insufficient information	–	8 (89)	–
Conflicting information	–	4 (45)	–
Other disorder	–	8 (89)	–
ASD ruled out	–	1 (11)	–

[a] Case versus non-case

[b] Chi-square goodness-of-fit test *p* values

[c] Documentation of a previous ASD diagnosis

[d] Child eligible for special education services

[e] Child eligible for autism special education services

[f] EERs were requested to explain reason for selecting Suspected case or DNQ as final case status