

**Evidence of a reduction over time in the behavioral severity of Autistic Disorder
diagnoses**

Andrew J.O. Whitehouse Ph.D.,¹ Matthew N. Cooper B Sc.,¹ Keely Bebbington B Sc.,¹ Gail
Alvares PhD.,¹ Ashleigh Lin Ph.D., John Wray MD,² Emma J. Glasson Ph.D.¹

Institutional affiliations: ¹Telethon Kids Institute, The University of Western Australia, 100
Roberts Rd, Subiaco 6009, Western Australia, Australia; ²Child Development Service, WA
Department of Health, 4-16 Rheola Street, West Perth 6005, Western Australia, Australia

Running title: Reduction over time in the severity of autism

Number of text pages: 12

Number of tables: Three

Number of figures: Three

Number of supplementary tables: Four

Number of supplementary figures: One

Address correspondence to: Prof Andrew Whitehouse, Telethon Kids Institute, University
of Western Australia, 100 Roberts Road, Subiaco, Western Australia 6008, Email:
Andrew.Whitehouse@telethonkids.org.au, Phone: +61 8 9489 7770, Facsimile: +61 8 9489
7700.

Funding Source: National Health and Medical Research Council; Grant number:
APP1077966 and APP1072593.

Lay abstract

The rise in the number of people diagnosed with Autism Spectrum Disorders (ASD) may in part be due to an increase in the diagnosis of individuals with less severe behaviours. The current study examined whether there were changes over time in the behavioural severity of individuals who received the diagnosis of Autistic Disorder. Data were from a register of new ASD diagnoses in Western Australia (n = 1252). From 2000 to 2006, we examined differences in the both percentage of newly diagnosed individuals meeting each of the 12 criterion for Autistic Disorder as well as severity ratings of the behaviours observed (not met, partially met, mild/moderate and extreme). We found a statistically significant reduction from 2000 to 2006 in the percentage of new diagnoses meeting two of 12 criteria. There was also a reduction across the study period in the proportion of new cases rated as having extreme severity on six criteria. There was a reduction in the proportion of individuals with three or more criteria rated as extreme from 2000 (16.0%) to 2006 (1.6%), while percentage of new cases with no 'extreme' rating on any criteria increased from 58.5% to 86.6% across the same period. This study provides the first clear evidence of a reduction over time in the behavioral severity of individuals diagnosed with Autistic Disorder. A shift towards diagnosing individuals with less severe behavioral symptoms may have contributed to the increasing prevalence of Autistic Disorder diagnoses.

Scientific abstract

The increasing prevalence of Autism Spectrum Disorders (ASD) may in part be due to a shift in the diagnostic threshold that has led to individuals with a less severe behavioural phenotype receiving a clinical diagnosis. The current study examined whether there were changes over time in the qualitative and quantitative phenotype of individuals who received the diagnosis of Autistic Disorder. Data were from a prospective register of new diagnoses in Western Australia (n=1252). From 2000 to 2006, we examined differences in both the percentage of newly diagnosed cases that met each criterion as well as severity ratings of the behaviours observed (not met, partially met, mild/moderate and extreme). Linear regression determined there was a statistically significant reduction from 2000 to 2006 in the percentage of new diagnoses meeting two of 12 criteria. There was also a reduction across the study period in the proportion of new cases rated as having extreme severity on six criteria. There was a reduction in the proportion of individuals with three or more criteria rated as extreme from 2000 (16.0%) to 2006 (1.6%), while percentage of new cases with no 'extreme' rating on any criteria increased from 58.5% to 86.6% across the same period. This study provides the first clear evidence of a reduction over time in the behavioral severity of individuals diagnosed with Autistic Disorder during a period of stability in diagnostic criteria. A shift towards diagnosing individuals with less severe behavioral symptoms may have contributed to the increasing prevalence of Autistic Disorder diagnoses.

Key words: Autistic Disorder, Autism Spectrum Disorder, epidemiology, diagnosis, prevalence

Introduction

Autism Spectrum Disorder (ASD) is the present day diagnostic term for neurodevelopmental disabilities characterized by impairments in social interaction, verbal and nonverbal communication and by repetitive patterns of behavior (American Psychiatric Association, 2013). In the absence of any consistent biological markers, a diagnosis of ASD is based on an appraisal of the qualitative nature and quantitative severity of these behavioral symptoms. ASD was first recognised in the third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III, American Psychiatric Association, 1980) as Infantile Autism, which described a highly constrained behavioral phenotype characterised by relatively severe symptomatology. Subsequent research identified that individuals who are less severely affected in terms of language ability and IQ may also present with the behavioural characteristics of ASD (Bishop, 1989; Georgiades et al. 2007), and more recent versions of the DSM incorporated flexibility into the behavioral requirements for a diagnosis. For example, DSM-IV (American Psychiatric Association, 1994) introduced the category of Autistic Disorder, also included in the text revision (American Psychiatric Association, 2000), which listed 12 criteria across three behavioural domains: Social impairments, communication difficulties, and restricted and repetitive behaviors and interests (RRBI). Individuals were deemed to meet the diagnostic threshold if they demonstrated specified combinations of six or more behavioural criteria across the three domains.

There is evidence that the broadening of diagnostic criteria from DSM-III to DSM-IV has played a role in the increasing numbers of individuals diagnosed with an ASD over the past three decades (Bishop, Whitehouse, Watt & Line, 2008; King & Bearman, 2009; Leonard et al., 2010; Miller et al., 2012; Shattuck, 2006). In a systematic review of prevalence studies,

Williams and colleagues found that epidemiological studies using DSM-IV criteria led to prevalence estimates three times those of studies using previous diagnostic criteria (Williams, Higgins, & Brayne, 2006). However, there is also evidence that the prevalence of ASD has continued to rise beyond the immediate period after the introduction of DSM-IV (Boyle et al., 2011; Maenner & Durkin, 2010; Parner et al., 2011). A recent population-based study in Sweden reported a significant increase in the prevalence of Autistic Disorder diagnoses between the years 1993 and 2002, but no rise in the level of autistic symptoms in the general population (Lundström et al., 2015). These findings raise the possibility that the increasing prevalence of ASD may have been driven by a shift in the diagnostic threshold over time towards diagnosing individuals with less severe autistic behaviours. However, no study has directly tested this hypothesis or quantified the effect.

The current study examined data from a register of newly diagnosed cases of ASD in Western Australia. Data are collected prospectively and include information on the diagnostic criteria met by each new case, as well as the severity of behaviours relevant to each criterion. We investigated all new cases that received the diagnosis of Autistic Disorder (diagnosed using DSM-IV-TR guidelines) from 2000 to 2006. The number of cases of Autistic Disorder in Western Australia has increased steadily over this period (Parner et al., 2011). The current study investigated the stability of criterion severity ratings over this seven-year time period during which the DSM-IV-TR criteria were stable, and tested the hypothesis that there has been population-level changes in the behavioural severity of individuals receiving a diagnosis of Autistic Disorder.

Methods

Cohort

Western Australia has a population of approximately 2.6 million people (10% of Australia's population). The majority (77%) of Western Australia's population live in the capital city, Perth (Australian Bureau of Statistics, 2011), and the vast majority of ASD diagnostic and intervention services are coordinated from this centre. Since 1997, Western Australia has had a standardized procedure for diagnosis of ASDs and eligibility for Government-supported therapy provision (Glasson et al., 2008). For children younger than 12 years of age, the diagnostic assessment must be performed by a team of three independent health professionals (paediatrician or psychiatrist, psychologist, and speech-language pathologist), and for adolescents (12-17 years) and adults, an assessment must be carried out by a clinical psychologist, a paediatrician (for adolescents) and/or a psychiatrist, with the inclusion of a formalised assessment by a speech-language pathologist as needed.

The Western Australian Register for Autism Spectrum Disorders (hereafter, 'Autism Register') is an independent, prospective collection of demographic and diagnostic data established January 1999 for newly diagnosed cases of ASDs in Western Australia (Glasson, 2002). At the time of diagnosis, the diagnosing clinicians submit information to the Autism Register using a dedicated form. Notification of new cases to the Autism Register is voluntary, but ascertainment during the period covered by this study (2000 – 2006 inclusive) was enhanced by annual archival reviews of the four Government centres at which the majority of diagnostic assessments in Western Australia were conducted. One of these centres was also the primary administrative contact in Western Australia for families enrolled in Government supported ASD therapy services, which are available to all children in this

jurisdiction with a diagnosis of Autistic Disorder. The diagnostic information of any cases found not to have been notified to the Autism Register was added to the annual collections, but without identifying details attached. Data collection has ethical approval from the Princess Margaret Hospital Human Research Ethics Committee (294EP).

In order to compare phenotypic changes over time in a large number of individuals diagnosed under consistent criteria, the current study included those cases who were diagnosed with Autistic Disorder (under DSM-IV-TR guidelines) in Western Australia between the years 2000 and 2006 (inclusive). These years were selected for analysis because during this period: (1) diagnostic assessments were predominantly conducted through four Government centres, and there was a central administrative portal for archival reviews, which provided confidence in achieving near-complete ascertainment of new cases of Autistic Disorder; (2) diagnosticians were strongly encouraged to complete severity scores for each criterion, and a near complete data-set was available; and (3) numerous studies (Boyle et al., 2011; Maenner & Durkin, 2010), including those in Western Australia (Parner et al., 2011), have found that this period corresponded to a particular increase in the prevalence of Autistic Disorder.

There were 1303 individuals who met the study inclusion criteria. Of these cases, 823 were notified to the Autism Register and 480 were added through archival reviews. A small number of cases ($n = 51$) were excluded from analyses because complete data on the criteria met and/or severity ratings were not provided by the diagnosing clinician(s), which left a final sample size of 1252 cases.

Diagnostic and behavioral characteristics

At the time of diagnosis, diagnosticians completed a data collection form and submitted this to the Autism Register. The form included DSM-IV-TR diagnostic criteria and diagnosticians were asked to rate the criteria for each new diagnostic case according to a four point scale: (0) Criteria definitely not met; (1) Criteria questionable/partially met, (2) Clearly meets criterion to mild or moderate degree, and (3) Clearly meets criterion to an extreme degree. Diagnosticians also indicated if a criterion was not assessed.

Diagnosticians indicated on the data collection form whether the Vineland Adaptive Behavior Scales–Version 1 (VABS; Sparrow, Balla, & Cicchetti, 1984) was completed at the time of diagnosis. Those who completed the VABS were asked to provide standardized scores for the Adaptive Behavior Composite Score and the Communication, Daily Living Skills, Socialisation and Motor Skills domains (if completed). Lower standardized scores represent greater functional impairment.

Statistical analysis

Data consisted of one record per individual. Multiple variables were created for each diagnostic criterion including: (a) the original rating (0 to 3 scale), (b) a dichotomous criteria met or not met (met being defined as a score of 2 or 3) variable; and (c) a dichotomous rating of ‘extreme’ or ‘not extreme’ (extreme being defined as a score of 3) variable. These were aggregated to counts and percentage per calendar year of diagnosis.

All data analyses were completed using *R* (R Core Team, 2012). Analyses included investigations of: (a) the age and sex distribution of the sample across the study period; (b) change in percentage of cases deemed to meet diagnosis criteria (individually); (c) change in percentage of individuals deemed to be displaying extreme behaviors for diagnosis criteria

(individually and aggregated), and (d) the number of criteria diagnosed as extreme (per case). Percentage change across calendar years within the study period was quantified using linear regression. As a sensitivity analysis, unit record data were analysed under a generalised linear model, modelled under the binomial distribution with the log link function, to calculate risk ratios. These analyses included adjustment for age and sex. It is likely that individuals who receive an ASD diagnoses during adulthood have different behavioral profiles to individuals who are diagnosed during childhood. To investigate the robustness of using whole population data, the analyses above were replicated restricting the data to only those individuals who were 18 years of age or younger at the time of diagnosis. Interaction terms between sex and calendar year were included, with models compared via ANOVA to assess the significance of the interaction term.

VABS standard scores were available for less than half of the current sample (see Supplementary Table 3 for numbers), which is insufficient ascertainment to determine whole population trends over time. However, these scores provided an opportunity to examine the validity of clinician severity ratings. Separate linear regression analyses were conducted to examine the relationship between VABS standardized scores and the relevant DSM-IV-TR criteria severity ratings. While the associations between VABS scores and autistic symptomatology in the ASD population are often weak, the highest correlations are observed between scores on the VABS Socialization domain and autistic behaviours in the social domain of DSM criteria (Klin et al., 2007; Perry, Flanagan, Dunn Geier & Freeman, 2009), and so analyses focused on these relationships.

Results

Demographic data

For the 1252 individuals who met the inclusion criteria, there was no statistically significant association between age at diagnosis or sex and calendar year across the study period (Table 1). The mean age of diagnosis was 6.47 (SD = 4.72) years and 82.5% of the sample was male.

Changes in diagnostic criteria met

Linear regression identified a statistically significant decrease in the number of criteria deemed as 'met' per individual across the study period, decreasing by -0.067 (95% CI; -0.108,-0.026; $p < 0.01$) per calendar year (Table 2). However this was not deemed clinically significant as the mean count was 8.22 in 2000 and only dropped to 8.02 in 2006, with the lowest value observed in 2005 (7.87).

Table 2 also presents the annual count and percentage of individuals meeting each DSM-IV-TR criterion. There was a statistically significant decrease per calendar year in the percentage of new diagnostic cases meeting criteria 1b (*failure to develop peer relationships*), -2.5% (95% CI -4.3, -0.8; $p = .013$), and 3a (*preoccupation with restricted interests*), -2.0% (95% CI -3.0,-1.0; $p = .003$). There was also a borderline significant ($p < 0.1$) decrease in the percentage of new diagnostic cases meeting criteria 2a (*delay in, or total lack of, spoken language*), -1.4% (95% CI -3.0, 0.2; $p = .08$), and 3b (*inflexible adherence to routines*), -1.8% (95% CI -3.7, 0.1; $p = .057$). There was a significant increase per calendar year in new diagnostic cases meeting criteria 3d (*persistent preoccupation with parts of an object*), 2.3%

(95% CI 1.2, 3.5; $p = .004$). No statistically significant changes over time were observed for the remaining criteria.

Changes in behavioral severity of diagnostic criteria met

Figure 1 presents change over time in the proportion of individuals who were rated as meeting a given criterion to an 'extreme' degree. Statistically significant decreases were observed for criteria 1a, 1b, 1c, 2b, 3a and 3b (all p values $< .05$), with the smallest statistically significant rate of decrease observed for criteria 2b (-0.7% per year, 95% CI -1.4, -0.1; $p = .031$) and the largest observed for criteria 1a (-2.1% per year, 95% CI -3.1,-1.2; $p = .002$). There was also a significant decrease per year in the percentage of individuals with any diagnostic criteria rated as 'extreme' (-3.8% per year, 95% CI -5.4,-2.2; $p = 0.002$, Figure 2). Table 3 presents beta coefficients from linear regression analyses for each criterion as well as across the three domains. The adjusted R-squared values suggest a large portion of the variation present is explained by calendar year. A sensitivity analysis was carried out modelling individual risk of receiving a severity rating of 'extreme' for each criterion over time, adjusting for age and sex. The magnitude and direction of the results (risk ratios) were in line with those above and are presented in Supplementary Table 1a. The risk ratio estimates presented in Table 3 saw no major deviation when these analyses were repeated excluding the 20 (1.6%) of subjects aged over 18 years at the time of diagnosis, presented in Supplementary Table 1b. The interaction terms between sex and calendar year were not significant, and Supplementary Table 2 and Supplementary Figure 1 show the consistency in observations and trend estimates between males and females.

The percentage of individuals with 'extreme' ratings for at least three criteria declined from 16.0% in 2000 to 1.6% in 2006, while the percentage of new cases with no 'extreme' ratings

increased from 58.5% to 86.6% across the same period (Figure 3a). There was no increase across calendar year observed in the number of individuals presenting with an 'extreme' diagnosis for at least one criteria, -1.3 (95% CI -3.6, 1.1; $p = .218$), which contrasts with a statistically significant increase per year in the number of individuals with no extreme ratings, 21.7 per calendar year (95% CI 14.0, 29.4; $p = .001$) (Figure 3b).

Validity of clinician severity ratings: VABS

Supplementary Table 3 presents VABS standard scores as a function of clinician severity ratings. Numerous main effects were identified, including for Adaptive Behavior Composite scores in relation to severity ratings on Criteria 1b and 1c, for the Communication subscale scores in relation to ratings on Criteria 1a and 1c, and for the Socialization subscale scores in relation to ratings on Criteria 1b and 1c. Linear regression also revealed a significant effect of calendar year on the VABS Adaptive Behavior Composite score and Socialization subscale, indicating increasing VABS scores (less functional impairment) from 2000 to 2006 (Supplementary Table 4).

Discussion

The current study provides the first evidence for population-level changes over time in the qualitative and quantitative phenotype of Autistic Disorder. Across the study period (2000 to 2006), there was a statistically significant reduction in the proportion of new diagnostic cases meeting two criteria (1b and 3a), and a reduction in the proportion of new cases rated as having 'extreme' severity for 6 of the 12 diagnostic criteria (1a, 1b, 1c, 2b, 3a and 3b). We found a significant reduction in the proportion of individuals given an extreme rating on three or more criteria from 2000 (16.0%) to 2006 (1.6%), while the percentage of new cases with

no 'extreme' rating on any criteria increased from 58.5% to 86.6% across the same period. The observed changes in the diagnostic phenotype are even more striking given that the same criteria (DSM-IV-TR) were used during the entire study period, and the jurisdiction of Western Australia had a rigorous and standardized protocol for Autistic Disorder diagnoses.

There was a statistically significant reduction over time in the endorsement of criteria 1b (*failure to develop peer relationships to appropriate developmental level*) and a borderline significant decrease in the endorsement of criteria 2a (*delay in, or total of, the development of spoken language*). These criteria are reflective of broader developmental abilities, and the observed changes appear to indicate a reduction over time in the proportion of new diagnoses with significant functional difficulties. There was also a statistically significant reduction over time in the endorsement of 3a (*preoccupation with restricted interests*) and a borderline significant reduction in the endorsement of 3b (*inflexible adherence to routines/rituals*), but an increase in the proportion of new diagnoses meeting criterion 3d (*persistent preoccupation with parts of an object*). It is arguable that the behaviours consistent with criteria 3a and 3b represent greater functional impairment to an individual than those reflected by criterion 3d, and that this qualitative shift in diagnostic characteristics is also indicative of a reduction over time in the proportion of new diagnoses with a high degree of functional difficulties.

The major strength of the study design is the prospective, population-level data from a jurisdiction that had a standardised and rigorous diagnostic protocol for Autistic Disorder (Glasson, 2002). Despite notification to the Autism Registry being voluntary, we have reason to believe that the current sample represents near-complete ascertainment of all diagnoses of 'Autistic Disorder' in Western Australia from 2000 to 2006. During the study period, the vast majority of diagnostic assessments for Autistic Disorder were made at four centres, including

one centre which was the primary administrative contact for Government-supported therapeutic services (eligible to all individuals with a diagnosis of Autistic Disorder). Annual archival reviews were conducted at each of these centres to identify and add to the Autism Register any cases that were not initially notified. We cannot rule out the possibility that there were a small number of cases who received both privately-funded diagnostic and therapeutic services and were not notified to the Autism Register. However, even cases diagnosed at small private clinics operating in Western Australia between 2000 and 2006 were required to be registered centrally for service provision and these pathways were approached for any missed cases. For this reason, we believe that the cohort of individuals who received a diagnosis of Autistic Disorder but were not added to the Autism Register, if any, would represent a small number of cases with less severe phenotypes, and would not alter the main conclusions of the study.

Clinician ratings of the severity of the behaviours relevant to each criterion was another strength of the study design. While severity ratings are part of the diagnostic classifications for ASD under DSM-5 guidelines, the use of ratings was not standard practice for diagnoses under DSM-IV-TR in other jurisdictions. It is important to highlight that the severity ratings had broad definitions, and that inter-rater reliability for clinician ratings could not be established from existing data. However, the VABS was completed on a proportion of cases, and we were able to demonstrate that standard scores on this measure (particularly the Total Composite, and the Communication and Socialization subscales) varied as a function of clinician severity ratings.

The current findings may inform our understanding of why the prevalence of ASD diagnoses has increased markedly over the past 50 years from 0.05% in the 1960s (Lotter, 1966) to at

least 1% in the 2010s (Elsabbagh et al., 2012). In the current study, we found that the number of individuals with no diagnostic criteria rated as 'extreme' in severity increased by an average of 21 cases per year from 2000 to 2006, whereas there was no statistically significant increase in the number of individuals with at least one diagnostic criterion rated as 'extreme'. This observation indicates that the majority of the increase in Autistic Disorder diagnoses in Western Australia during this period (Parner et al., 2011) is likely due to an increase in the diagnoses of individuals with less severe behavioural phenotypes. A range of sociological factors may have contributed to changes over time in the behavioural profiles of individuals presenting for a diagnostic assessment, including greater awareness of ASD among clinicians and the community (Charman 2002; Leonard et al., 2010) and the introduction of Government funded services specific to individuals with a diagnosis of Autistic Disorder (Gurney et al., 2003; Leonard et al., 2010; Mandell & Palmer, 2005; Nassar et al., 2009).

The current findings also have implications for future revisions of the diagnostic criteria for ASD. The population-level changes observed in the current study were across a study period that commenced six years after the introduction of the diagnostic category of Autistic Disorder (in DSM-IV in 1994). The findings indicate that population-level changes to the behavioural phenotype of community-based diagnoses can occur even during periods of stability in diagnostic criteria. The DSM-5 included several significant changes to the DSM-IV-TR diagnostic criteria for ASD, including combining the social and communication behavioral domains into a single 'social communication' domain. There are significantly fewer combinations of symptoms that lead to an ASD diagnosis under DSM-5 guidelines, which several studies have indicated may decrease the number of individuals diagnosed with ASD (Bennet & Goodall, 2016; Kulage, Smaldone, & Cohn, 2014). While the longer-term effects of the new diagnostic criteria on the behavioral phenotype of ASD is uncertain,

clinician severity ratings are a key part of the DSM-5 criteria for ASD, which will facilitate ongoing monitoring.

Acknowledgements

The authors would like to acknowledge the Advisory Committee of the Western Australian Register for Autism Spectrum Disorders, as well as the diagnosticians and participants who contributed data to the Register over a long period.

References

- American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: American Psychiatric Association.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. revised). Washington, DC: American Psychiatric Association.
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: APA Press.
- Australian Bureau of Statistics (2011). Greater Perth (Greater Capital City Statistical Area). Census QuickStats.
- Bennett, M., & Goodall E. (2016). A meta-analysis of DSM-5 autism diagnoses in relation to DSM-IV and DSM-IV-TR. *Review Journal of Autism and Developmental Disorders*, 3, 119-124.
- Bishop, D. V. M., Whitehouse, A. J. O., Watt, H. J., & Line, E. A. (2008). Autism and diagnostic substitution: evidence from a study of adults with a history of developmental language disorder. *Developmental Medicine and Child Neurology*, 50(5), 341-345.
- Bishop, D.V.M. (1989). Autism, Asperger's syndrome and semantic-pragmatic disorder: Where are the boundaries? *British Journal of Disorders of Communication*, 24, 107-121.
- Boyle, C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M., ... Kogan, M. D. (2011). Trends in the prevalence of developmental disabilities in US children, 1997–2008. *Pediatrics*.
- Charman, T. (2002). The prevalence of autism spectrum disorders. *European Child & Adolescent Psychiatry*, 11(6), 249-256. doi: 10.1007/s00787-002-0297-8

- Elsabbagh, M., Divan, G., Koh, Y.-J., Kim, Y. S., Kauchali, S., Marcín, C., ... Fombonne, E. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism Research*, 5(3), 160-179.
- Georgiades, S., Szatmari, P., Zwaigenbaum, L., Duku, E., Bryson, S., Roberts, W., ... Mahoney, W. (2007). Structure of the autism symptom phenotype: A proposed multidimensional model. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46(2), 188-196.
- Glasson, E.J., MacDermott, S., Dixon, G., Cook, H., Chauvel, P., Maley-Berg, A., Wray., J. (2008). Management of assessments and diagnosis for children with autism spectrum disorders: The Western Australian model. *Medical Journal of Australia*, 2008, 188, 288–91.
- Glasson, E. J. (2002). The Western Australian Register for Autism Spectrum Disorders. *Journal of Paediatrics and Child Health*, 38(3), 321-321.
- Gurney, J. G., Fritz, M. S., Ness, K. K., Sievers, P., Newschaffer, C. J., & Shapiro, E. G. (2003). Analysis of prevalence trends of autism spectrum disorder in minnesota. *JAMA Pediatrics*, 157(7), 622-627.
- King, M., & Bearman, P. (2009). Diagnostic change and the increased prevalence of autism. *International Journal of Epidemiology*, 38(5), 1224-1234.
- Klin, A., Saulnier, C. A., Sparrow, S. S., Cicchetti, D. V., Volkmar, F. R., & Lord, C. (2007). Social and communication abilities and disabilities in higher functioning Individuals with autism spectrum disorders: The Vineland and the ADOS. *Journal of Autism and Developmental Disorders*, 37(4), 748-759.
- Kulage, K. M., Smaldone, A. M., & Cohn, E. G. (2014). How will DSM-5 affect autism diagnosis? A systematic literature review and meta-analysis. *Journal of Autism and Developmental Disorders*, 44(8), 1918-1932.

- Leonard, H., Dixon, G., Whitehouse, A. J. O., Bourke, J., Aiberti, K., Nassar, N., ... Glasson, E. J. (2010). Unpacking the complex nature of the autism epidemic. *Research in Autism Spectrum Disorders*, 4(4), 548-554.
- Lotter, V. (1966). Epidemiology of autistic conditions in young children. *Social Psychiatry*, 1, 124–137.
- Lundström, S., Reichenberg, A., Anckarsäter, H., Lichtenstein, P., & Gillberg, C. (2015). Autism phenotype versus registered diagnosis in Swedish children: prevalence trends over 10 years in general population samples. *British Medical Journal*, 350, 1961.
- Maenner, M. J., & Durkin, M. S. (2010). Trends in the prevalence of autism on the basis of special education data. *Pediatrics*, 126(5), e1018-e1025.
- Mandell, D. S., & Palmer, R. (2005). Differences among states in the identification of autistic spectrum disorders. *Archives of Pediatrics & Adolescent Medicine*, 159(3), 266-269.
- Miller, J. S., Bilder, D., Farley, M., Coon, H., Pinborough-Zimmerman, J., Jenson, W., McMahon, W. M. (2012). Autism Spectrum Disorder reclassified: A second look at the 1980s Utah/UCLA autism epidemiologic study. *Journal of Autism and Developmental Disorders*, 43(1), 200-210.
- Nassar, N., Dixon, G., Bourke, J., Bower, C., Glasson, E., de Klerk, N., & Leonard, H. (2009). Autism spectrum disorders in young children: effect of changes in diagnostic practices. *International Journal of Epidemiology*, 38(5), 1245-1254. doi: 10.1093/ije/dyp260
- Parner, E. T., Thorsen, P., Dixon, G., de Klerk, N., Leonard, H., Nassar, N., ... Glasson, E. J. (2011). A comparison of autism prevalence trends in Denmark and Western Australia. *Journal of Autism and Developmental Disorders*, 41(12), 1601-1608.

Perry, A., Flanagan, H. E., Dunn Geier, J., & Freeman, N. L. (2009). The Vineland Adaptive Behavior Scales in young children with autism spectrum disorders at different cognitive levels. *Journal of Autism and Developmental Disorders*, 39(7), 1066-1078.

R Core Team. (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Shattuck PT. (2006). The contribution of diagnostic substitution to the growing administrative prevalence of autism in US special education. *Pediatrics*, 117(4), 1028–1037.

Sparrow, S., Balla, D., & Cicchetti D. (1984). *The Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Services.

Williams, J. G., Higgins, J. P. T., & Brayne, C. E. G. (2006). Systematic review of prevalence studies of autism spectrum disorders. *Archives of Disease in Childhood*, 91(1), 8-15.

Table 1. Participant demographics by calendar year

	2000	2001	2002	2003	2004	2005	2006	Total
N	106	152	167	178	192	203	254	1252
<i>Age</i>								
Mean (SD)	6.43 (5.47)	6.57 (4.75)	6.19 (4.37)	6.62 (4.68)	6.56 (4.41)	6.38 (5.15)	6.53 (4.51)	6.47 (4.72)
Range	2 - 43	2 - 25	1 - 26	1 - 33	2 - 26	1 - 38	1 - 31	1 - 43
<i>Sex*</i>								
Male	84 (79.2%)	122 (80.3%)	129 (77.2%)	152 (85.9%)	157 (82.2%)	171 (85.1%)	207 (84.5%)	1022 (82.5%)
Female	22 (20.8%)	30 (19.7%)	38 (22.8%)	25 (14.1%)	34 (17.8%)	30 (14.9%)	38 (15.5%)	217 (17.5%)

*Data on sex of case were missing for 1 case in each of 2003 and 2004, 2 cases in 2005 and 9 cases in 2006.

Table 2. The mean (SD) number of criteria endorsed by new Autistic Disorder diagnoses between 2000 and 2006. The number (%) of new cases meeting each diagnostic criterion across the study period are also presented.

Criteria	Assessment	2000	2001	2002	2003	2004	2005	2006
Number of criteria endorsed	M (SD)	8.22 (1.39)	8.55 (1.53)	7.98 (1.43)	8.04 (1.42)	8.21 (1.36)	7.87 (1.40)	8.02 (1.48)
1a	Met	89 (84.0)	128 (84.2)	135 (80.8)	154 (86.5)	167 (87.0)	165 (81.3)	211 (83.1)
Marked impairment in the use of multiple nonverbal behaviors	Not Met	17 (16.0)	24 (15.8)	32 (19.2)	24 (13.5)	24 (12.5)	38 (18.7)	43 (16.9)
	Not Assessed	0 (-)	0 (-)	0 (-)	0 (0.0)	1 (0.5)	0 (0.0)	0 (0.0)
1b	Met	89 (84.0)	135 (88.8)	135 (80.8)	149 (83.7)	154 (80.2)	145 (71.4)	183 (72.0)
Failure to develop peer relationships to appropriate developmental level	Not Met	16 (15.1)	17 (11.2)	30 (18.0)	27 (15.2)	34 (17.7)	42 (20.7)	57 (22.4)
	Not Assessed	1 (0.9)	0 (-)	2 (1.2)	2 (1.1)	4 (2.1)	16 (7.9)	14 (5.5)
1c	Met	80 (75.5)	112 (73.7)	117 (70.1)	130 (73.0)	145 (75.5)	144 (70.9)	177 (69.7)
Lack of spontaneous seeking to share enjoyment or interests with others	Not Met	26 (24.5)	40 (26.3)	49 (29.3)	48 (27.0)	47 (24.5)	59 (29.1)	77 (30.3)
	Not Assessed	0 (-)	0 (-)	1 (0.6)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
1d	Met	91 (85.8)	135 (88.8)	140 (83.8)	144 (80.9)	171 (89.1)	177 (87.2)	217 (85.4)
Lack of social or emotional reciprocity	Not Met	15 (14.2)	17 (11.2)	27 (16.2)	34 (19.1)	21 (10.9)	26 (12.8)	36 (14.2)
	Not Assessed	0 (-)	0 (-)	0 (-)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.4)
2a	Met	84 (79.2)	108 (71.1)	123 (73.7)	123 (69.1)	144 (75.0)	136 (67.0)	174 (68.5)
Delay in, or total lack of, the development of spoken language	Not Met	22 (20.8)	42 (27.6)	43 (25.7)	55 (30.9)	47 (24.5)	66 (32.5)	79 (31.1)
	Not Assessed	0 (-)	2 (1.3)	1 (0.6)	0 (0.0)	1 (0.5)	1 (0.5)	1 (0.4)
2b*	Met	46 (43.4)	79 (52.0)	75 (44.9)	93 (52.2)	101 (52.6)	100 (49.3)	134 (52.8)
Impairment in the ability to initiate or sustain a conversation	Not Met	9 (8.5)	17 (11.2)	17 (10.2)	12 (6.7)	11 (5.7)	14 (6.9)	31 (12.2)
	Not Assessed	51 (48.1)	56 (36.8)	75 (44.9)	73 (41.0)	80 (41.7)	89 (43.8)	89 (35.0)
2c	Met	44 (41.5)	81 (53.3)	75 (44.9)	59 (33.1)	72 (37.5)	78 (38.4)	112 (44.1)

Stereotyped and repetitive use of language or idiosyncratic language	Not Met	37 (34.9)	46 (30.3)	52 (31.1)	77 (43.3)	79 (1.1)	78 (38.4)	101 (39.8)
	Not Assessed	25 (23.6)	25 (16.4)	40 (24.0)	42 (23.6)	41 (21.4)	47 (23.2)	41 (16.1)
2d	Met	88 (83.0)	135 (88.8)	134 (80.2)	153 (86.0)	157 (81.8)	162 (79.8)	206 (81.1)
Lack of varied, spontaneous, make-believe play or social imitative play	Not Met	18 (17.0)	17 (11.2)	32 (19.2)	25 (14.0)	32 (16.7)	38 (18.7)	46 (18.1)
	Not Assessed	0 (-)	0 (-)	1 (0.6)	0 (0.0)	3 (1.6)	3 (1.5)	2 (0.8)
3a	Met	71 (67.0)	105 (69.1)	109 (65.3)	115 (64.6)	112 (58.3)	118 (58.1)	149 (58.2)
Preoccupation with restricted interests abnormal in intensity or focus	Not Met	35 (33.0)	47 (30.9)	57 (34.1)	63 (35.4)	79 (41.1)	84 (41.4)	102 (39.8)
	Not Assessed	0 (-)	0 (-)	1 (0.6)	0 (0.0)	1 (0.5)	1 (0.5)	5 (0.8)
3b	Met	64 (60.4)	88 (57.9)	87 (52.1)	84 (47.2)	90 (46.9)	98 (48.3)	131 (51.6)
Inflexible adherence to specific, non-functional routines or rituals	Not Met	42 (39.6)	64 (42.1)	79 (47.3)	94 (52.8)	102 (53.1)	105 (51.7)	122 (48.0)
	Not Assessed	0 (-)	0 (-)	1 (0.6)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.4)
3c	Met	52 (49.1)	82 (53.9)	81 (48.5)	99 (55.6)	109 (56.8)	107 (52.7)	135 (53.1)
Stereotyped and repetitive motor mannerisms	Not Met	53 (5-)	70 (46.1)	85 (50.9)	79 (44.4)	81 (42.2)	95 (46.8)	119 (46.9)
	Not Assessed	1 (0.9)	0 (-)	1 (0.6)	0 (0.0)	2 (1.0)	1 (0.5)	0 (0.0)
3d	Met	73 (68.9)	112 (73.7)	122 (73.1)	128 (71.9)	154 (80.2)	167 (82.3)	210 (82.7)
Persistent preoccupation with parts of an object	Not Met	33 (31.1)	40 (26.3)	44 (26.3)	50 (28.1)	37 (19.3)	36 (17.7)	43 (16.9)
	Not Assessed	0 (-)	0 (-)	1 (0.6)	0 (0.0)	1 (0.5)	0 (0.0)	1 (0.4)

*only completed for those individuals with 'adequate speech'

Table 3. Beta coefficients for calendar year from linear regression modelling of the proportion of individuals rated as ‘extreme’ on each criterion (2000 to 2006).

Criteria	Beta (95%CI)	P value	Adjusted R ²
Social domain			
Criteria 1a	-2.12 (-3.09, -1.15)	p < 0.01	0.84
Criteria 1b	-1.81 (-2.69, -0.94)	p < 0.01	0.82
Criteria 1c	-1.51 (-2.05, -0.98)	p < 0.01	0.90
Criteria 1d	-0.86 (-1.73, 0.01)	p = 0.05	0.47
Social combined	-2.79 (-3.82, -1.77)	p < 0.01	0.89
Communication			
Criteria 2a	-1.40 (-3.34, 0.53)	p = 0.12	0.29
Criteria 2b	-0.73 (-1.36, -0.10)	p = 0.03	0.57
Criteria 2c	-0.27 (-0.73, 0.20)	p = 0.20	0.17
Criteria 2d	-1.0 (-2.51, 0.51)	p = 0.15	0.24
Communication combined	-2.32 (-4.57, -0.07)	p = 0.05	0.50
RRBI domain			
Criteria 3a	-1.38 (-2.55, -0.21)	p = 0.03	0.58
Criteria 3b	-1.14 (-2.08, -0.20)	p = 0.03	0.59
Criteria 3c	-0.47 (-0.99, 0.05)	p = 0.07	0.42
Criteria 3d	-0.39 (-0.91, 0.13)	p = 0.11	0.31
RRBI combined	-2.48 (-3.98, -0.99)	p = 0.01	0.74
All criteria combined	-3.78 (-5.38, -2.19)	p < 0.01	0.86

Figure captions

Figure 1. The proportion of Autistic Disorder diagnoses in each calendar year that had an ‘extreme’ score on a given DSM-IV-TR criterion. Criteria are grouped by domain, and lines of best fit are included for the trend across years. Horizontal jitter used to separate overlapping data points.

Figure 2. The proportion of Autistic Disorder diagnoses in each calendar year that had an ‘extreme’ score on any criteria in a given domain. Lines of best fit are included for the trend across years. Horizontal jitter used to separate overlapping data points.

Figure 3. Stacked bar plot showing the percentage (Figure 3a) and count (Figure 3b) of Autistic Disorder diagnoses in each calendar year who are rated as having ‘extreme’ behavior on DSM-IV-TR criteria.

Supplementary Figure 1. The proportion of Autistic Disorder diagnoses in each calendar year that had an ‘extreme’ score on a given DSM-IV-TR criterion, separated for males and females. Horizontal jitter used to separate overlapping data points.