Short Report: A National E-Cohort Study of Incidence and Prevalence of Autism

Underwood, Jack F G[1]; DelPozo-Banos[2], Marcos; Frizzati[3], Aura; John, Ann[2]; Hall, Jeremy[1]

1. Neuroscience and Mental Health Research Institute, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK
2. Institute of Life Sciences, Medical School, Swansea University, Swansea, UK
3. Cedar Healthcare Technology Research Centre, Cardiff & Vale University Health Board, Cardiff, UK

Lay Abstract

Autism Spectrum Disorders (Autism) are thought to be relatively common, with analyses estimating 1-2% in the population could meet diagnostic criteria. There is debate over whether the number of new cases is rising, with some suggesting an increase, whilst other studies finding no change. In this study we take anonymised healthcare record data from more than 3.6 million people in Wales to produce a national estimate of prevalence and incidence. We found the overall prevalence rate of autism in the population was 0.51%. The number of new-recorded cases of autism increased from 0.19% per 1000 person-years in 2001 to 0.64% per 1000 person-years in 2016. The estimate of 0.51% prevalence in the population is lower than suggested by previous studies, potentially due to underdiagnosis or lack of recognition by services. From 2000-2016 the number of autism services for adults has increased, and autism is more widely known in society. We suggest that these extra services and awareness have contributed to the 3.4-fold rise in new autism diagnoses over the study and the >150% increase in the years 2008-2016. Our study suggests that while the number of people being diagnosed with autism is increasing, many are still missed by healthcare services.

Introduction

Autism Spectrum Disorders (hereafter Autism) are a group of complex neurodevelopmental disorders characterised by persistent difficulties in social interaction and communication, stereotypic, repetitive and restricted behaviours and interests, and difficulty coping with change (World Health Organisation, 2016). Autism is common, with epidemiological studies reporting a prevalence of 1-2% in the population, with a male to female ratio of around 3:1 (Lai et al., 2014). Estimates of the prevalence in children suggest it has been rising since the 1980s, with Centre for Disease Control and Prevention surveillance reports documenting an increase from approximately one in 150 children during 2000-2002, to approximately one in 68 during 2010-2012 (Redfield et al., 2014). However, data from the UK General Practice Research Database (GPRD) suggests that autism incidence in children is static (Taylor et al., 2013). Earlier identification and assistance for children with autism enables adaptation and support, and can improve social communication and reduce anxiety (Lai et al., 2014). As autism spectrum conditions are lifelong neurodevelopmental conditions, underpinned by

atypical neural system development, it is anticipated that increases in childhood diagnosis will correspond to increased rates amongst adults, though this has not yet been substantiated (Lai et al., 2014). The aims of this study were to determine the prevalence and incidence of autism within the Welsh population. Validation of incidence data on a population level offers the opportunity to inform future healthcare policy and provision, leading to improvements in quality of life for adults with autism.

Methodology

Data sources, Population and Settings

This was a retrospective population based electronic cohort study. Data was sourced from the Secure Anonymised Information Linkage (SAIL) Databank (www.saildatabank.com) (Lyons et al., 2009). SAIL is a data repository of anonymised person based linkable data from healthcare and public settings. Data access was approved by the SAIL independent Information Governance Review Panel in 2018 (application 0843). The policies, permissions, processes, structures and controls applied to manage the SAIL databank are described elsewhere (John et al., 2018).

This study used the Welsh Longitudinal General Practice dataset (WLGP) and PEDW (Patient Episode Database for Wales) for extracting relevant diagnostic data, and the Welsh Demographic Service Dataset (WDSD), which includes anonymised demographics, geographic index (LSOA = Lower Layer Super Output Area index 2011) and deprivation data (Welsh Index of Multiple Deprivation quintile 2011). The WDSD is an administrative register of all individuals in Wales to use NHS services, including anonymised demographics and general practice registration.

3.63 million individuals from the Welsh population over the period 2000-2016 were included in the study. The PEDW dataset comprises hospital admissions (inpatient and day case) clinical information for all NHS Wales hospitals. At the time of the analysis 333 (of 432, 77%) general practices in Wales were supplying their data to SAIL for the WLGP dataset, covering 79% of the population. Individuals with autism were identified in the WLGP and PEDW datasets using ICD (International Classification of Diseases) version 10 codes and Read Codes v2. This methodology has been validated in previous studies of other psychiatric disorders (Lloyd et al., 2015). The autism cohort utilised ICD-10 codes: F840, F841, F845, plus the 'legacy' codes F84.., F842, F843, F844, F848, and F849. The Read v2 Codes used were: 1J9.., E140., E1400, E1401, E140z, Eu840, Eu841, Eu845, E141., Eu84., Eu842, Eu843, Eu844, Eu846, Eu84y, and Eu84z.

Statistical Analysis

Data in the SAIL databank were interrogated using structured query language (SQL DB2). Analysis of outputs were processed utilising IBM SPSS Statistics 23 for Window (IBM Corp. Released, 2015). For analysis of incidence and prevalence, the denominator included all individuals alive and living in Wales and registered with a GP providing data to SAIL between 1st of January 2000 and 31st of December 2016. Age during this period or at time of diagnosis

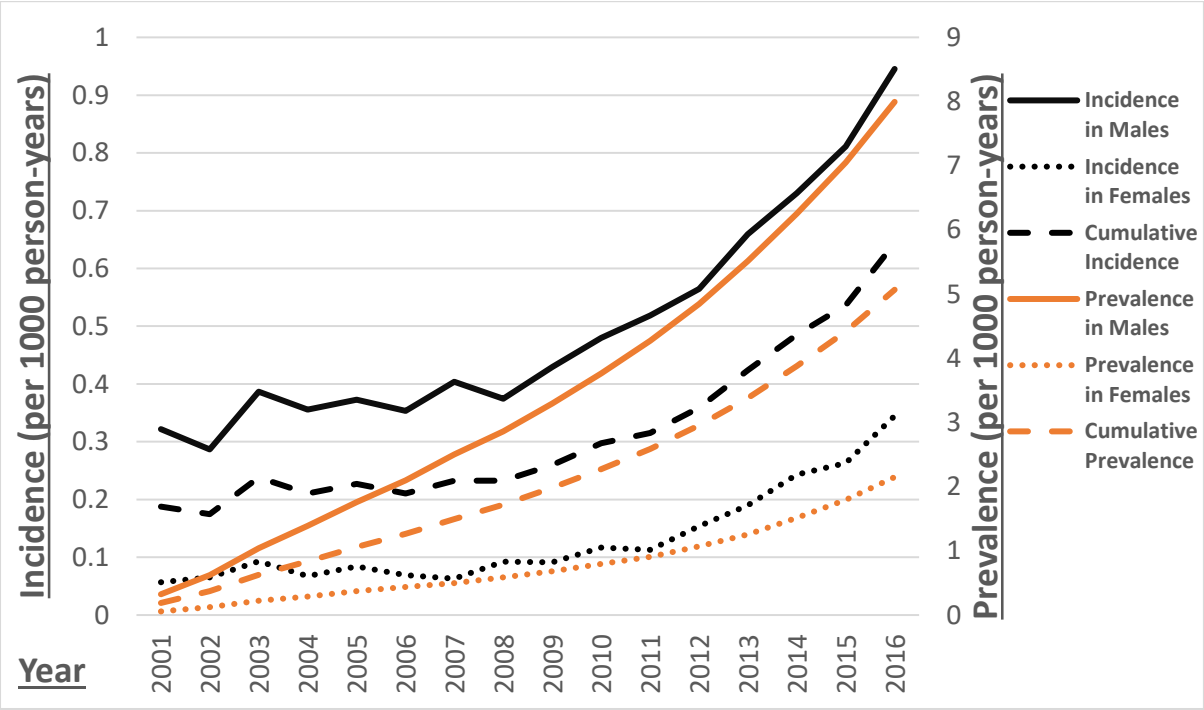
was not relevant. Prevalence of autism and annual rates of first ever incident recorded autism were calculated for the whole 16-year period or until death, whichever was earlier. A new incident episode was defined as an entry in the records with no previous entry of that problem ever recorded and with at least six months' worth of WLGP data prior to that diagnosis. Annual first ever incident rates were calculated per person years at risk. Person-time at risk was calculated using the start of each year (1st January) or start of registration (plus six months), whichever later. End date was earliest of: date of leaving practice supplying SAIL, date of death, or end of year (31st December).

Results

Autism Incidence and Prevalence

The yearly incidence for new recording in health record of autism diagnosis incorporating all 3,632,078 SAIL participants increased over the study period (2001-2016), from 0.19% per 1000 person-years in 2001, up to a maximum of 0.64% per 1000 person-years in 2016 (Figure 1), a 3.4-fold increase. This change comprised an increase in incidence in both male and female individuals, from 0.32% and 0.06% per 1000 person-years at risk up to 0.95% and 0.35% per 1000 person-years respectively for 2016 (Figure 1).

Figure 1: Incidence and Prevalence of Autism in the SAIL dataset over assessed years (2001-2016)



Incidence and Prevalence calculated utilising person years at risk (see Methodology). Cumulative incorporates male and female data.

The prevalence of autism also increased across the assessed period, from 0.19 per 1000 people in 2001 to 5.07 per 1000 people in 2016 (Figure 1). This cumulative figure summarises the rise in prevalence seen amongst both male and female participants. Male

prevalence rose from 0.32 per 1000 people in 2001 to 7.99 per 1000 people in 2016. Female prevalence rose from 0.06 per 1000 people in 2001 to 2.15 per 1000 people in 2016.

Discussion

In this large population-based electronic-cohort study we draw on anonymised data from General Practice and Hospital Patient Encounters for over 3.6 million individuals representative of the population of Wales to produce national estimates of autism prevalence and incidence. We found a prevalence of autism of 0.51% in 2016, and a 3.4-fold increase in autism incidence over the study period from 0.19% per 1000 person-years in 2001, up to a maximum of 0.64% per 1000 person-years in 2016.

Estimates point to a prevalence of autism in the general population around 1-2% (Brugha et al., 2011; Lai et al., 2014; Lord et al., 2020; Redfield et al., 2014; Russell et al., 2015; Simonoff et al., 2008; Taylor et al., 2013). Prior studies of autism prevalence have examined a number of birth cohorts, CDC surveillance data on children's health records in the US, and the General Practice Research Database (GPRD) for general practice health records of children in the UK (Elsabbagh et al., 2012; Redfield et al., 2014; Russell et al., 2015; Taylor et al., 2013). Brugha *et al* were reportedly the first to ascertain data for adults, finding a prevalence of 0.98-1.1% using a national stratified survey methodology (Brugha et al., 2011, 2016). Our prevalence estimate of 0.51% is lower than these epidemiological studies, which may be due to methodological differences reflecting the under-identification and reporting of autism in healthcare records.

We demonstrate a marked increase in incidence of autism diagnosis, particularly in the eight-year period from 2008-2016. During this time new recorded diagnoses increased by >150%. Despite this, the absolute incidence rate remains low. Multiple groups have reported that rates of autism diagnosis are increasing over time, though with a more recent plateau (Lai et al., 2014; Russell et al., 2015). Comparative data from the General Practice Research Database (GPRD) has demonstrated that the incidence rate amongst children 2-8 years during an overlapping time period (2001-2010) remained static (Taylor et al., 2013). We, like others, would suggest that this increase is due to recognition, increased societal awareness, changes in practice and improved availability of diagnostic services particularly for adults with autism (Lai et al., 2014; Lord et al., 2020; Russell et al., 2015).

This study utilised the SAIL Database, a nationally-representative, high quality anonymised e-cohort of >3.6 million individuals (Lyons et al., 2009). We reviewed General Practice and Hospital Encounter datasets, comprising ICD-10 and Read v2 Codes for diagnoses under analysis. As an e-cohort study, exposure and outcome classification errors cannot be excluded as we are unable to verify the diagnoses. Electronic records of General Practice data prior to 2000 are sparsely held in the SAIL Databank, and therefore prevalent cases prior to 2000 may be missing from this estimate. This explains the dramatic increase from very low baseline of our reported prevalence. We rely upon accurate documentation by healthcare staff, and just 21.67% of our sample had an autism diagnosis in both datasets. It is likely our study misses individuals with autism in the community who have not come into contact with their GP or Hospital services resulting in their diagnosis not being on record.

Our study is therefore likely to be an underestimation. Diagnostic overshadowing, the practice of one diagnosis taking primacy and leading to underdiagnosis of co-occurring conditions, is also recognised to be a problem within autism (Brugha, 2018). Both of these limitations may compound to generate errors in our estimates and contribute to lower figures, although our large sample size and replication of comparable results would suggest these errors are minimal. We have attempted to compensate through use of both primary and secondary care data, and these limitations are common to e-cohort and record linkage studies which depend upon individuals' attendance, recognition and recording to be accurate.

In conclusion, in our national population sample we found a marked rise in incidence over the study period to 2016. This is consistent with improved societal recognition and access to diagnostic services resulting in greater rates of diagnosis; incident cases not picked up in childhood being diagnosed as adults. Despite this we found lower prevalence rates than demonstrated in previous studies, suggesting that a large proportion of those adults with autism are not recognised by the healthcare system.

Funding

This project was funded by a Wellcome Trust ISSF Clinical Primer award, and subsequent Wellcome Trust GW4-CAT Clinical Doctoral Fellowship to JFGU. JH and JFGU are supported by Healthcare Research Wales through the National Centre for Mental Health (NCMH) and an MRC Pathfinder Grant (MC_PC_17212). MDPB and AJ are supported by Healthcare Research Wales (grants CA04 and SC35296), and an MRC Pathfinder Grant (MC_PC_17211).

Contributors

JFGU principally designed the study, acted as principal investigator, statistically analysed the data and wrote the first draft of the manuscript with input from AF. AF, MDPB, AJ and JH all contributed to the design of the study. AF and MDPB co-assisted with data retrieval and analysis. All authors critically read the manuscript and contributed to the data interpretation and writing. All authors approved the final version of the manuscript.

Declaration of interests

JH has received grants from Takeda for research work unrelated to this project. The other authors declare no competing interests.

References

Brugha, T. S. (2018). Anxiety and depression in adults with autism spectrum disorder: a systematic review and meta-analysis. *Psychological Medicine*. https://doi.org/10.1017/S0033291718002283

Brugha, T. S., McManus, S., Bankart, J., Scott, F., Purdon, S., Smith, J., Bebbington, P., Jenkins, R., Meltzer, H., H, M., AM, R., CE, R., D, S., & GC, W. (2011). Epidemiology of Autism Spectrum Disorders in Adults in the Community in England. *Archives of General Psychiatry*, *68*(5), 459. https://doi.org/10.1001/archgenpsychiatry.2011.38

Brugha, T. S., Spiers, N., Bankart, J., Cooper, S. A., McManus, S., Scott, F. J., Smith, J., & Tyrer, F. (2016). Epidemiology of autism in adults across age groups and ability levels. *British Journal of*

Psychiatry, *209*(6), 498–503. https://doi.org/10.1192/bjp.bp.115.174649

Elsabbagh, M., Divan, G., Koh, Y. J., Kim, Y. S., Kauchali, S., Marcín, C., Montiel-Nava, C., Patel, V., Paula, C. S., Wang, C., Yasamy, M. T., & Fombonne, E. (2012). Global Prevalence of Autism and Other Pervasive Developmental Disorders. *Autism Research*, *5*(3), 160–179. https://doi.org/10.1002/aur.239

IBM Corp. Released. (2015). IBM SPSS Statistics for Windows, Version 23.0. *2015*.

John, A., McGregor, J., Jones, I., Lee, S. C., Walters, J. T. R., Owen, M. J., O'Donovan, M., DelPozo-Banos, M., Berridge, D., & Lloyd, K. (2018). Premature mortality among people with severe mental illness — New evidence from linked primary care data. *Schizophrenia Research*. https://doi.org/10.1016/J.SCHRES.2018.04.009

Lai, M. C., Lombardo, M. V, & Baron-Cohen, S. (2014). Autism. *The Lancet*, *383*(9920), 896–910. https://doi.org/10.1016/S0140-6736(13)61539-1

Lloyd, K., McGregor, J., John, A., Craddock, N., Walters, J. T., Linden, D., Jones, I., Bentall, R., Lyons, R. A., Ford, D. V., & Owen, M. J. (2015). A national population-based e-cohort of people with psychosis (PsyCymru) linking prospectively ascertained phenotypically rich and genetic data to routinely collected records: Overview, recruitment and linkage. *Schizophrenia Research*, *166*(1–3), 131–136. https://doi.org/10.1016/j.schres.2015.05.036

Lord, C., Traolach, S. B., Tony, C., James, C., Guillaume, D., Thomas, F., Emily J. H., J., Rebecca M., J., Andrew, P., Matthew W., S., Julie, L. T., & Jeremy, V.-V. (2020). Autism Spectrum Disorder. *Nature Reviews*, *6*(5), 1–23. https://doi.org/10.1038/s41572-019-0138-4

Lyons, R. A., Jones, K. H., John, G., Brooks, C. J., Verplancke, J.-P., Ford, D. V, Brown, G., & Leake, K. (2009). The SAIL databank: linking multiple health and social care datasets. *BMC Medical Informatics and Decision Making*, *9*(1), 3. https://doi.org/10.1186/1472-6947-9-3

Redfield, R. R., Kent, C. K., Leahy, M. A., Martinroe, J. C., Spriggs, S. R., Yang, T., Doan, Q. M., King, P. H., Maitland, P. D., Starr, T. M., Yang, M., Jones, T. F., Boulton, M. L., Caine, V. A., Daniel, K. L., Fielding, J. E., Fleming, D. W., Halperin, W. E., Holmes, K. K., … Schaffner, W. (2014). *Morbidity and Mortality Weekly Report Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years-Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014 Centers for Disease Control and Prevention MMWR Editorial and Production Staff (Serials) MMWR Editorial Board*.

Russell, G., Collishaw, S., Golding, J., Kelly, S. E., & Ford, T. (2015). *Changes in diagnosis rates and behavioural traits of autism spectrum disorder over time*. https://doi.org/10.1192/bjpo.bp.115.000976

Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric Disorders in Children With Autism Spectrum Disorders: Prevalence, Comorbidity, and Associated Factors in a Population-Derived Sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(8), 921–929. https://doi.org/10.1097/CHI.0b013e318179964f

Taylor, B., Jick, H., & MacLaughlin, D. (2013). Prevalence and incidence rates of autism in the UK: Time trend from 2004-2010 in children aged 8 years. *BMJ Open*, *3*(10). https://doi.org/10.1136/bmjopen-2013-003219

World Health Organisation. (2016). ICD-10 Version:2016. In *Who*. https://doi.org/10.1177/1071100715600286