MENTAL HEALTH

Cohort effects explain the increase in autism diagnosis among children born from 1992 to 2003 in California

Katherine M Keyes,1,2 Ezra Susser,1,2 Keely Cheslack-Postava,1,3 Christine Fountain,3 Kayuet Liu3 and Peter S Bearman3*

1Department of Epidemiology, Columbia University, New York, NY, USA, 2New York State Psychiatric Institute, New York, NY, USA and 3Paul F. Lazarsfeld Center for the Social Sciences, Columbia University, New York, NY, USA

*Corresponding author. Cole Professor of the Social Sciences, Institute for Social and Economic Research and Policy, Columbia University, 420 W. 118th Street, IAB Building, 814, New York, NY 10027, USA. E-mail: psb17@columbia.edu

Accepted 1 November 2011

Background The incidence and prevalence of autism have dramatically increased over the last 20 years. Decomposition of autism incidence rates into age, period and cohort effects disentangle underlying domains of causal factors linked to time trends. We estimate an age-period-cohort effect model for autism diagnostic incidence overall and by level of functioning.

Methods Data are drawn from sequential cohorts of all 6,501,262 individuals born in California from 1992 to 2003. Autism diagnoses from 1994 to 2005 were ascertained from the California Department of Development Services Client Development and Evaluation Report.

Results Compared with those born in 1992, each successively younger cohort has significantly higher odds of an autism diagnosis than the previous cohort, controlling for age and period effects. For example, individuals born in 2003 have 16.6 times the odds of an autism diagnosis compared with those born in 1992 [95% confidence interval (CI) 7.8–35.3]. The cohort effect observed in these data is stronger for high than for low-functioning children with an autism diagnosis.

Discussion Autism incidence in California exhibits a robust and linear positive cohort effect that is stronger among high-functioning children with an autism diagnosis. This finding indicates that the primary drivers of the increases in autism diagnoses must be factors that: (i) have increased linearly year-to-year; (ii) aggregate in birth cohorts; and (iii) are stronger among children with higher levels of functioning.

Keywords Autism, cohort effects, time trends, USA

Introduction The prevalence of autism has increased in many countries across the last two decades yet the reasons for the increase remain highly controversial. Whereas a large part of the increase remains unexplained, four main hypotheses have been supported by empirical data as contributing to the increase in prevalence:

(i) diagnostic definition (e.g. publication of DSM-IV);
(ii) diagnostic accretion (i.e. children with an initial diagnosis of mental retardation acquiring an autism diagnosis) and expansion (i.e. children with autism at the higher end of the functioning spectrum being diagnosed with greater...
frequency);\textsuperscript{13,15} (iii) increased awareness of signs and symptoms of autism;\textsuperscript{17} and (iv) individual-level risk factors that have increased in frequency (e.g. older parental age at birth\textsuperscript{19}).

Age-period-cohort analyses of data collected over time can disentangle the time-varying forces shaping trends over time and help to focus investigations of underlying aetiology, as each effect implicates a broad domain of causal factors. In general, age-period-cohort models decompose variance in trends over time into those attributable to age-, period- and cohort-effects. At the individual level, autism diagnosis is strongly related to child age,\textsuperscript{19,20} and substantial evidence indicates that the average age of diagnosis has decreased in more recently born cohorts.\textsuperscript{21} Age effects (i.e. changes in the age structure of the population or age at diagnostic ascertainment), however, are unlikely to be driving the increased incidence of autism diagnosis as age-adjusted rates still show a marked increase over time.\textsuperscript{22} Cohort effects can be conceived of as changes in health status that are confined to or stronger among particular age groups in particular time periods.\textsuperscript{23} Cohort effects could arise in these data through a number of potential mechanisms. For example, if the prevalence of a risk factor acting at conception exhibits change over time (e.g. paternal age), then each successively younger cohort will be differentially exposed, manifesting as a cohort effect (e.g. each successively younger cohort has older fathers). Alternatively, an exposure introduced into the population as a whole could differentially affect autism incidence depending on age of exposure. This would also manifest as a cohort effect. In contrast, if period effects are observed to be the main driver of increased autism diagnosis, then risk factors which have also varied across time but have similar effects across all age groups may be considered as potentially implicated in the increase in autism. Period effects can be conceived of as changes in disease status that affect all age groups simultaneously, and often coincide with widespread environmental or diagnostic nosological changes but may also reflect widespread societal changes in some circumstance. The key in differentiating period from cohort effects is that if period effects are operative, then the incidence of autism should increase across all age groups under study at a particular time, rather than among specific age groups at a particular time.

Evidence to date is strongly suggestive of powerful cohort effects in the incidence and prevalence of autism.\textsuperscript{21,24,25} For example, data from the Centres for Disease Control and Prevention funded Autism and Developmental Disabilities Monitoring (ADDM) Network indicate a $>$2-fold increase in the prevalence of autism spectrum disorders among 8-year-olds in the US born in 1998 compared with 1994 (Morbidity and Mortality Weekly Report\textsuperscript{26}) (with substantial variation in the size of the increase across states\textsuperscript{24}) and more recent data indicate that the prevalence continues to increase.\textsuperscript{2} Population-based data from the California Department of Developmental Services (DDS) and Denmark indicate consistent increases in the incidence of autism diagnoses across age for each successively younger birth cohort from approximately 1990 forward.\textsuperscript{21,22} This previous work has not incorporated formal age-period-cohort modelling. Thus a quantification of the contribution of age, period and cohort effects in population-based surveillance data on autism trends is called for in order to tease apart the effect of cohort from other time-related trends that vary by period and age.

Finally, previous studies have not examined age-period-cohort effects in autism diagnostic incidence by level of function at time of diagnosis. Some have suggested that increases in autism in recent decades have been primarily restricted to high-functioning children with autism.\textsuperscript{27} Functioning is not a diagnostic designation; here we define functioning in terms of a child’s reported ability in the domains of communication and social interaction. If expanded inclusion of high-functioning children in the diagnostic pool is entirely responsible for changes in the incidence of autism diagnosis, then any observed age-period-cohort trends should be restricted to higher-functioning children with autism. Thus, examining age-period-cohort trends in both low-functioning and high-functioning children with autism provides a robust and innovative method to examine effects of diagnostic expansion.

The present study utilized a population-registry of all California births and all diagnoses of autism registered with the California DDS in order to comprehensively characterize age, period and cohort effects in autism diagnostic incidence from 1994 to 2005, covering cohorts born from 1992 to 2003. Our main aim was to analyse age-period-cohort trends in the diagnosis of autism. Additionally, we examined age-period-cohort models for low- and high-functioning children with autism.

**Methods**

**Data source**

We utilize data on all 6,501,262 individuals born in California from 1992 to 2003, the cohorts for which we currently have information available. Data were drawn from California birth records as well as Client Development and Evaluation Report (CDER) records from the DDS, an agency coordinating diagnoses, services and support for individuals with developmental disabilities living in California. We matched DDS records on 21,093 individuals with an ICD-9 diagnosis of autism born in California between 1992 and 2003 to their California birth records using probabilistic and deterministic matching algorithms based on first, middle and last name, birth date, race, zip code at birth and sex. Further information...
on matching procedures and details of the data can be found elsewhere.\textsuperscript{13,19} Incidence rate per 10 000 person-years was determined by creating person-time for each respondent: individuals were at risk for a diagnosis each year from age 2 to age 12-years (diagnoses before age 2 or after age 12 are rare\textsuperscript{19}) until an autism diagnosis was received. Four individuals received a diagnosis before age 2 years and were excluded. Thus, each individual in the data set had a maximum of 10 points of data, one for each year in which they were at risk of receiving an autism diagnosis. Once an individual received an autism diagnosis, they no longer contributed person-time. Subsequent evaluations were not considered in this analysis. The final analysis included 43 693 205 time points.

**Measures**

**Autism diagnosis**

In California the DDS system is responsible for coordinating diagnoses, services and support for persons with developmental disabilities including autism. The DDS provides services to patients with full syndrome autism, but not to those with other spectrum disorders or pervasive developmental disorders unless they have another qualifying condition or substantial disability. The vast majority of persons with autism in California are enrolled with the DDS, making it the largest administrative source of data on autism diagnoses.\textsuperscript{14} Autism diagnoses were ascertained by extraction of DDS records for all clients with a CDER on file between 1994 and 2005. Trained diagnosticians screened each potential client upon entry into the system and CDER records were updated after annual evaluations.

**Functioning**

As in our previous studies,\textsuperscript{17} we derived a ‘functioning’ score from a global index of function on two dimensions relevant to autism: social interaction, and communication and language, both recorded on the CDER at the time of intake. These scores were created from evaluations by developmental specialists based on observation and caregiver reports. These evaluations are designed to allocate services, and in the process of the evaluation, developmental specialists record scores on structured assessment forms regarding social interaction, communication and language. Social interaction functioning was based on five items [internal consistency–reliability (\(\alpha\)) from 0.85 to 0.86 by birth cohort], and communication functioning was based on three items (\(\alpha\) from 0.77 to 0.84 by birth cohort). These dimensions were equally weighted and normalized within age groups. We created cutpoints based on upper and lower quintiles of the distributions within our data. For the purpose of the present study, those above the age-standardized 80th percentile (using the 1992 cohort deciles as the referent) were considered high-functioning, those in the 20–79th percentile were considered mid-functioning and those below the 20th percentile were considered low-functioning. Note that this is relative to other included subjects. As such, ‘high-’ and ‘low-’ functioning are neither diagnostic definitions nor were they designations used explicitly in the allocation of services. Sensitivity analyses using a range of alternative cutpoints did not change the results.

**Statistical methods**

**Graphical analysis**

We began our assessment of age, period and cohort effects by thorough examination of incidence rates by graphical analysis. This informs not only the overall assessment of age, period and cohort effects, but also the type of and specifications for statistical modelling.

**Statistical model**

The best practice for statistical modelling of age, period and cohort effects has been a source of scientific and biostatistical debate for over four decades.\textsuperscript{23,28–30} Briefly, because age, period and cohort are linearly related (Cohort = Period – Age), simultaneous consideration of the linear effects of the three variables in a statistical model results in a non-identified regression matrix in a least squares framework. When considering all possible age-period-cohort methods, the constraint-based approach\textsuperscript{31} required minimally restrictive assumptions and provided the best theoretical fit for this research question. We constrained the model such that the incidence of autism was constant after age 8 years, based on data indicating this to be the case (see Figure 1). Birth cohort and period were unconstrained.

Our constraint-based age-period-cohort model was a generalized linear model with a logit link function, in which autism diagnosis was the outcome and age, period and cohort were categorical predictors. All models used generalized estimating equations to adjust standard errors for repeated measures (multiple years for each child) with an auto-regressive correlation matrix specification. We tested each of the final models (overall, and by functioning) for goodness-of-fit using the Barnhart & Williamson approach.\textsuperscript{32} Sensitivity analyses were run changing the reference groups in all categories to determine whether results were dependent on specific reference groups used; no changes in interpretation were detected.

**Results**

**Graphical description of age, period and cohort trends**

Age-, period- and cohort-specific incidence rates per 10 000 person-years of observation are shown in Figure 1. As data were sparse for the 2003 cohort, we present results for the 1992 to 2002 cohorts only. By age, autism diagnostic incidence peaks at age 3 (11 per 10 000 person-years), decreases
throughout early childhood and remains stably low after age 7 (1 per 10,000 person-years). By period, no appreciable trend emerges; the rate is relatively constant at 4 per 10,000 person-years, save for a change to 3 per 10,000 person-years in 1999. After 1996, rates remain stable, save for increases from 4 per 10,000 person-years in 2001 to 5 per 100,000 person-years in 2002 and 2004. By cohort, the incidence of autism diagnosis consistently increases with each successively younger birth cohort.

The flat period graph reflects dual forces simultaneously operating; whereas the number of diagnoses substantially increases in the later years, reflecting the increasing rate among younger birth cohorts, the amount of person-time in the denominator also increases over time, as older age groups with individuals less likely to receive a new diagnosis continue to contribute person-time. In summary, the combination of increasing numbers of children with autism due to younger cohorts with increasing person-time due to older children still contributing person-time renders a relatively flat graph indicative of a lack of substantial period effects.

Figure 2 shows the age-specific incidence of autism diagnosis by birth cohort in California among those born 1992–2003.

Statistical age-period-cohort model

As expected, controlling for period and cohort, age remains strongly related to the odds of autism diagnosis (results shown in Supplementary Table S1);
for example, those aged 3 have 37 times the odds of diagnosis of those aged 2 years [95% confidence interval (CI) 31.7–43.9]. Results for period and cohort effects are shown in Table 1. We observe a significant effect of period; controlling for age and birth cohort, those observed in 1995 have 3.75 times the odds of an autism diagnosis compared with those observed in 1994 (95% CI 2.4–5.9); after 1995, the period effect in autism diagnoses is mostly stable. We observe evidence for a strong and mostly linear cohort effect that is large in magnitude and statistically significant. Compared with those born in 1992, each successively younger cohort has significantly higher odds of an autism diagnosis, and odds ratios (OR) increase monotonically across cohorts compared with the 1992 cohort.

Shown in Figure 3 is the predicted probability of an autism diagnosis based on the age-period-cohort model for three ages, periods and birth cohorts. This figure provides a visual representation of the contribution of all three effects to the incidence of autism diagnosis based on the model. By isolating each component, we can see the effect of birth cohort controlling for age and period; the effect of period controlling for age and birth cohort; etc. For example, the six columns aggregated under ‘Calendar Year 2005’ show the predicted probability of autism for 3- to 5-year-olds in the 1992 and 2005 birth cohorts, holding period at the 2005 estimate (whereas it is impossible for 3- to 5-year-olds born in 1992 or 2005 to be actually observed in 2005, we set it to the 2005 effect, thus, the conditions that were present in 2005 as predicted by the model).

A hypothetical group of 3-year-olds born in 1992 have a predicted incidence of 1.2 per 10,000 person-years, holding the estimate for period effect at its 2005 level. A hypothetical group of 3-year-olds born in 2000 have a predicted incidence of 6.8 per 10,000 person-years, again holding the period effect constant at its 2005 level. This is indicative of the strong cohort effect in these data.

### Discussion

We found evidence of a significant birth cohort effect in the incidence of autism diagnoses in California from 1994 to 2005 that is stronger for high-functioning children with autism. This finding indicates that the primary drivers of the increases in autism diagnoses are factors that: (i) have increased linearly year-to-year; (ii) aggregate in birth cohorts; and (iii) affect those with higher levels of functioning to a greater degree than those with lower levels of functioning. Although age-period-cohort models do not test specific hypotheses about underlying mechanisms through which trends arise, they do provide a useful lens through which to evaluate these hypotheses. To be consistent with our findings, a hypothesized mechanism should predict a linear increase over time in successive birth cohorts and a stronger time

### Table 1

<table>
<thead>
<tr>
<th>Birth cohort</th>
<th>OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1993</td>
<td>1.15</td>
<td>(1.0–1.3)</td>
</tr>
<tr>
<td>1994</td>
<td>1.38</td>
<td>(1.2–1.6)</td>
</tr>
<tr>
<td>1995</td>
<td>1.74</td>
<td>(1.4–2.1)</td>
</tr>
<tr>
<td>1996</td>
<td>2.15</td>
<td>(1.7–2.8)</td>
</tr>
<tr>
<td>1997</td>
<td>2.69</td>
<td>(2.0–3.7)</td>
</tr>
<tr>
<td>1998</td>
<td>3.32</td>
<td>(2.3–4.8)</td>
</tr>
<tr>
<td>1999</td>
<td>4.25</td>
<td>(2.8–6.6)</td>
</tr>
<tr>
<td>2000</td>
<td>5.97</td>
<td>(3.7–9.8)</td>
</tr>
<tr>
<td>2001</td>
<td>7.93</td>
<td>(4.6–13.8)</td>
</tr>
<tr>
<td>2002</td>
<td>12.07</td>
<td>(6.5–22.3)</td>
</tr>
<tr>
<td>2003</td>
<td>16.62</td>
<td>(7.8–35.3)</td>
</tr>
</tbody>
</table>

*Table based on a constraint-based GEE regression model simultaneously controlled for age, period and cohort categories.*
trend for high than for low-functioning children with an autism diagnosis.

For example, one hypothesis for the increased incidence of autism is that increasing social awareness is driving the trends. If increasing awareness of autism (or its impact on diagnosis) were increasing linearly over time, and were concentrated on children of a specific age (e.g. age 2 and 3 years), it would be
manifest as a cohort effect, similar to what we observed. In addition, to be consistent with our findings, increasing awareness must have been stronger for high than for low-functioning children with autism at that age. On the other hand, if increasing social awareness were gradually diffusing and being applied across the range of ages (mainly 2–8 years) when children are diagnosed with autism, it should have been manifest as a period effect. In each successive calendar year we would see an increase in diagnoses across this range of ages (i.e. each year, parents would be more likely to consider whether their 2- to 8-year-old children may have autism, and clinicians more likely to refer them to services for evaluation of autism diagnosis). However, this pattern is not consistent with our data. These alternative processes for increasing social awareness can be empirically tested, in order to better evaluate the degree to which increasing social awareness could be driving the time trend in autism diagnoses.

As a second example, the hypothesis that childhood vaccines are main drivers of increasing autism incidence could be examined through this lens. Vaccination aggregates by birth cohort. Vaccination patterns have not, however, increased linearly year-to-year in California. In our view, this makes it unlikely that vaccination could have produced the cohort effect observed here.

We note too that a hypothesized confluence of factors could be evaluated through this lens. Consider the confluence of increasing social awareness, changes in diagnostic practice and changes in incidence related to family structure (e.g. parental age at birth of first child has been increasing). If these factors are concentrated in children of a specific age, the confluence could produce a cohort effect. One could still observe a cohort effect if the different components within such a confluence of factors had a somewhat different impact at different time points (e.g. more complete ascertainment affecting increasing incidence in early cohorts, and diagnostic expansion affecting rates of diagnosis in later cohorts).

Although we observe evidence of a small period effect in our model, with an increase in population incidence across age coinciding with the publication of DSM-IV in 1994, it is likely that this is a methodological artefact. Due to the structure of the data, we only have information on 2-year-old children in 1994; in 1995, 3-year-olds are added to the data. As the incidence of autism diagnosis is low among 2-year-olds and relatively high among 3-year-olds, a heightened incidence of autism in 1995 in the population is likely a reflection of adding 3-year-olds to the data. Thus we caution against overinterpretation of the observed period effect from 1994 to 1995. The lack of period effects could reflect the relatively constrained age range over which autism diagnoses typically occur. Although changes in diagnostic practice could be plausible drivers of a potential period effect, these changes are likely to affect those in specific age groups more than others, in which case changes in diagnostic criteria would manifest as a cohort effect. Overall, given that environmental factors would need to have the same effect on autism diagnoses among children of all ages in order for a period effect to explain the increases in autism incidence, our findings indicating the lack of substantial period effects correspond to the developmental epidemiology of autism.

The present study is limited by a number of factors. Age-period-cohort models are often criticized because the additional constraint necessary in order to estimate a model renders results prone to bias. These criticisms are valid when age-period-cohort models are estimated without background information, prior hypotheses and close examination of the research guiding the choice and specification of the statistical model. For these data we carefully examined the graphical trends in the data, estimated models with varying assumptions and ultimately chose the modeling strategy that best corresponded to our theory about mechanisms in these data and the observable graphical trends with the minimal number of identifying assumptions. As in all statistical models, however, our assumptions are unverifiable. These results should also be reviewed with the limitations of California administrative data in mind. Whereas these data represent diagnoses from 21 regional centres with similar distributions of functioning scores in all centres, the distribution of functioning has changed over time as children diagnosed with autism are increasingly likely to have higher-functioning scores. Only children with diagnoses of autism or substantial functional impairment are served by the California system, thus children with autism spectrum disorder on the higher end of the functioning spectrum may be missed in these data. However, we note that the average levels of social and verbal functioning as reported at first assessment have substantially increased across cohorts in these data, as demonstrated in our results. Nevertheless, these results can only be generalized to trends over time in autism rather than autism spectrum disorders more generally. Other limitations include use of birth records, which may have coding errors. However, date of birth is not typically recorded inaccurately, and any errors are assumed to be non-differential in nature. Next, because these data are linked with birth certificate records, individuals who are born in California but receive an autism diagnosis in another state are misclassified as not having autism in these data. This misclassification would affect estimates to the extent that migration patterns out of California have changed over time in ways that are associated with autism diagnosis. Whereas out-migration has increased in California, the rich services provided in California to children with autism give little ground to suspect that out-migration
from California is related to autism diagnosis. Finally, we cannot separate changes in referral patterns and practices from changes in prevalence in these data, as only children with autism that came to the attention of the California DDS are included in these data. These results have significant implications for autism research. This research area has been particularly controversial, with a number of factors posited to affect the observed increase including changes in diagnostic practice, younger age at diagnosis and heightened awareness. Our data do not support younger age at diagnosis as the driver of increased incidence, but are consistent with changes in diagnostic practice and heightened awareness as potentially viable explanations for the increase in autism diagnoses. We observe a weak period effect compared with the strong and linear cohort effect, suggesting that broad environmental factors that have equal effects across age cannot explain the increase in autism diagnosis. Further research into possible factors that increase linearly over time by birth cohort and are particularly salient for high-functioning children with autism should be pursued in further autism research.

Supplementary Data
Supplementary Data are available at IJE online.

Acknowledgements
Financial support was provided by the National Institutes of Health Director’s Pioneer Award Program, part of the NIH Roadmap for Medical Research (grant 1 DP1 OD003635-01). We thank Soumya Mazumdar and Alix Winter for helpful comments on earlier versions of this article and the California Department of Developmental Services for support of the study.

Conflict of interest: None declared.

KEY MESSAGES
- Risk of an autism diagnosis increased in each successively younger birth cohort from 1992 to 2003 independently of age and period effects.
- No appreciable period effects are found for autism diagnoses from 1992 to 2003.
- Cohort effects are more pronounced among children who display higher social and language functioning.

References


