Annals of Epidemiology 24 (2014) 260–266

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Original article

Population attributable fractions for three perinatal risk factors for autism spectrum disorders, 2002 and 2008 autism and developmental disabilities monitoring network

Laura A. Schieve PhD [a,*], Lin H. Tian MD [a], Jon Baio EdS [a], Kristin Rankin PhD [b], Deborah Rosenberg PhD [b], Lisa Wiggins PhD [a], Matthew J. Maenner PhD [c], Marshalyn Yeargin-Allsopp MD [a], Maureen Durkin PhD [c], Catherine Rice PhD [a], Lydia King PhD [d], Russell S. Kirby PhD [e], Martha S. Wingate DrPH [f], Owen Devine PhD [a]

[a] National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, GA
[b] School of Public Health, University of Illinois, Chicago
[c] Waisman Center, University of Wisconsin-Madison
[d] Medical University of South Carolina, Charleston
[e] School of Public Health, University of South Florida, Tampa
[f] College of Public Health, University of Alabama-Birmingham

ARTICLE INFO

Article history:
Received 23 July 2013
Accepted 23 December 2013
Available online 15 January 2014

Keywords:
Autism
Birth weight
Cesarean section
Population
Premature birth
Risk factors

ABSTRACT

Purpose: Numerous studies establish associations between adverse perinatal outcomes/complications and autism spectrum disorder (ASD). There has been little assessment of population attributable fractions (PAFs).

Methods: We estimated average ASD PAFs for preterm birth (PTB), small for gestational age (SGA), and Cesarean delivery (CD) in a U.S. population. Average PAF methodology accounts for risk factor co-occurrence. ASD cases were singleton non-Hispanic white, non-Hispanic black, and Hispanic children born in 1994 ($n = 703$) or 2000 ($n = 1339$) who resided in 48 U.S. counties included within eight Autism and Developmental Disabilities Monitoring Network sites. Cases were matched on birth year, sex, and maternal county of residence, race-ethnicity, age, and education to 20 controls from U.S. natality files.

Results: For the 1994 cohort, average PAFs were 4.2%, 0.9%, and 7.9% for PTB, SGA, and CD, respectively. The summary PAF was 13.0% (1.7%–19.5%). For the 2000 cohort, average PAFs were 2.0%, 3.1%, and 6.7% for PTB, SGA, and CD, respectively, with a summary PAF of 11.8% (7.5%–15.9%).

Conclusions: Three perinatal risk factors notably contribute to ASD risk in a U.S. population. Because each factor represents multiple etiologic pathways, PAF estimates are best interpreted as the proportion of ASD attributable to having a suboptimal perinatal environment resulting in PTB, SGA, and/or CD.

Published by Elsevier Inc.

Introduction

Autism spectrum disorders (ASDs) are characterized by social and communication impairments and restricted, stereotyped patterns of behavior [1]. ASD prevalence in the U.S. [2–4] and other populations [5–9] has increased markedly and is currently estimated at 1%–2% [2–9]. Numerous genetic factors are implicated in the etiology of ASDs [10] and twin studies suggest high heritability [11]; the composite evidence supports gene-environment interactions. Research on nongenetic risk factors is evolving. Numerous studies document associations between ASD and various adverse perinatal outcomes and complications [12]. Limited studies suggest associations between ASD and more specific maternal exposures such as infections [13,14], medications [15–17], and environmental pollutants [18,19].

For most ASD risk factors, there has been no assessment of population attributable fractions (PAFs). We estimated PAFs for three perinatal risk factors, preterm birth (PTB), small for gestational age (SGA), and Cesarean delivery (CD) among U.S. children included in the Autism and Developmental Disabilities Monitoring (ADDM) Network and compared PAFs from the most recent ADDM surveillance year with those from an earlier time. We chose factors that were both relatively common (≥10% population prevalence) and thus could substantively contribute to the population ASD burden and established as ASD risk factors through multiple studies in a range of populations [12]. Nonetheless, each factor represents a

* Corresponding author. National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, MS E-86, 1600 Clifton Road, Atlanta, GA 30333. Tel.: +1 404 498 3888; fax: +1 404 498 3550.
E-mail address: LSchieve@cdc.gov (L.A. Schieve).

1047-2797/$ – see front matter Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.annepidem.2013.12.014

composite of multiple potential underlying etiologic mechanisms. Their PAF estimates are thus best interpreted as the proportion of ASD attributable to having a suboptimal perinatal environment resulting in PTB, SGA, and/or CD.

Methods

Study population

The ADDM Network is an ongoing ASD surveillance program among 8-year-old children residing in selected U.S. population–based sites. Thirteen sites participated in the 2002 ADDM surveillance year and 14 participated in the 2008 surveillance year.

For children meeting birth year and residence eligibility criteria, each ADDM site reviews special education records and medical records from providers who conduct developmental evaluations. Records with documentation of an ASD diagnosis or education classification or behavioral characteristics consistent with possible ASD are fully abstracted. Data abstracted include demographics, ASD and other disability diagnoses, behavioral descriptions from developmental evaluations, and intelligence quotient (IQ) score. Abstractions from different sources for the same child are concatenated. Trained clinicians review the composite abstractions using a standardized protocol based on the *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition Text Revision) to classify children as having or not having ASDs [1]. Sites link their final data for ASD cases to state natality files; across sites 70% of children are born in-state and match a birth record.

Study population, cases

Our sample selection strategy is outlined in the Appendix. We initially selected children classified as ASD cases in 2002 or 2008 from 13 sites that participated in ADDM both years. Because ADDM tracks children aged 8 years, these children were born in 1994 and 2000. We further selected children residing both at birth and during the surveillance year in counties included in ADDM sites' catchment areas in both 2002 and 2008. This narrowed our population, as the geographic boundaries changed for some sites. In addition, the birth residence restriction (which was necessary to ensure comparability with controls) meant that we pragmatically restricted our population to sites that included the maternal residence county indicator in their submitted ADDM-natality data set (three sites did not) and to children linked to their birth record. We further excluded two sites that did not provide other needed variables. These selection criteria, although not impacting internal validity, did narrow the generalizability. Nonetheless, our defined study population still included 48 counties from eight states. Because of subgroup sample size constraints, we further limited the population to singleton non-Hispanic white (NHW), non-Hispanic black (NHB), and Hispanic children ($n = 747$ and 1406 cases from 2002 and 2008, respectively).

During analysis, we excluded a small percentage of children (3% from 2002 and 1% from 2008) missing data on one or more study variables and a small percentage of children (3% from both 2002 and 2008) included in a final matching stratum with a low number of potential controls per case (see the following section). Our final analytic sample included 703 children from 2002 ADDM (1994 birth cohort) and 1339 children from 2008 ADDM (2000 cohort).

Study population, controls

Although sites link their ADDM and natality datafiles, the deidentified data they submit for the pooled data set include only ASD cases (i.e., unlinked births from sites' natality files are not provided). Thus, we selected controls from public-use 1994 and 2000 U.S. natality files. We could not discern which births within those files were subsequently identified as ADDM cases (and thus, already included in our sample). Given the relatively low ASD population prevalence, the overall probability of selecting a case as a control was low.

To carefully and efficiently consider confounders, we used a matched design. We matched each case to 20 controls from the same birth year on sex, maternal race-ethnicity (NHW, NHB, Hispanic), county of residence, age (<20, 20–29, 30–34, 35+ years), and education (high school or less, greater than high school) at birth. We selected a high number of controls because the PAF methodology combined with modeling methods used resulted in a loss of controls within certain strata.

Public-use natality files do not include the specific maternal residence county for county populations less than 100,000. Rather, a general "small-county" indicator is provided. Thus, cases with a maternal county population of 100,000 or higher were exactly matched to controls on maternal residence county, whereas cases born to mothers from small-population counties were matched on the general small-county indicator for the state.

Given both number and type of matching factors, our sample was subdivided into numerous matching strata, some with a small number of births. Thus, one study selection criterion was birth within a study-matching stratum including a minimum of 20 potential controls.

Even still, some included strata were small and there was a nonnegligible possibility that we inadvertently selected the case as one of the controls. We conducted sensitivity analyses to assess the impact of this possible problem (see the following section).

Perinatal factors

PTB, SGA, and CD were derived from natality file data. PTB was defined as gestational age less than 37 completed weeks. Gestational age was based on last menstrual period or clinical estimate when last menstrual period was missing. SGA was defined as birth weight for gestational age less than 10th percentile of sex-specific referent curves for U.S. singleton livebirths between 1999 and 2000 [20]. CD included both primary and secondary CD.

Statistical methods

We estimated the summary PAF of exposure to any one of the three risk factors (PTB, SGA, or CD), alone or in any combination with any of the other factors. This represents the maximum proportion of ASD cases attributable to this risk factor set. Then, we partitioned the summary PAF into unique *average PAFs* for each risk factor to estimate proportions of ASD cases attributable to each factor on average, while considering the interaction among factors and the dynamic nature of the risk factors in the population. Average PAFs account for co-occurrence among the three factors while adjusting for other potential confounders; thus, in general, average PAFs address the inherent overestimation that occurs when computing separate crude or adjusted PAFs [21].

The average PAF is a summary estimate that considers all possible sequences of eliminating risk factors in a defined risk factor set. This methodology requires estimation of multiple sequential PAFs. For example, one sequence for this analysis addresses the hypothetical question, what would be the impact on ASD prevalence if one could first eliminate PTB from the population, followed by SGA, followed by CD. One sequential PAF is estimated for each factor in this sequence and an average PAF for each risk factor is derived using the simple average of all sequential PAFs for that factor.