Copy number variation in the dosage-sensitive 16p11.2 interval accounts for only a small proportion of autism incidence: A systematic review and meta-analysis

Kyle M. Walsh, MPhil, and Michael B. Bracken, PhD, MPH

Purpose: Autism is one of the most heritable complex disorders, but the genetic etiology of autism spectrum disorders is unexplained in ~90% of cases. Highly penetrant microdeletions and microduplications of 16p11.2 contribute to the pathogenesis of autism spectrum disorder, but the extent to which these variants account for the total burden of idiopathic autism spectrum disorders has not been systematically investigated. **Methods:** A systematic literature review and meta-analysis were performed to determine the prevalence of these variants among individuals diagnosed with autism spectrum disorders. A planned subgroup analysis was conducted to assess prevalence differences between sporadic and familial autism spectrum disorder cases. **Results:** In the combined analysis of 3613 idiopathic autism spectrum disorder cases from seven studies, the meta-analytic prevalence of these microdeletions and microduplications was 0.76% (95% CI, 0.51–1.12%). When stratified by copy number variant-type, the prevalence of microdeletions was 0.50% (95% CI, 0.31–0.82%) and the prevalence of microduplications was 0.28% (95% CI, 0.14–0.56%). Sporadic autism spectrum disorder cases showed only a slightly higher prevalence than familial cases. **Conclusion:** The number needed to test to identify one such variant is 132 patients (95% CI, 89–198). Such information, especially as it pertains to diagnostic yield in genetic testing, should prove useful to clinicians considering chromosomal microarray analysis in subjects with autism spectrum disorders. ***Genet Med* 2011:13(5):377–384.**

Key Words: *autism spectrum disorder, copy number variation, prevalence, meta-analysis, diagnostic yield*

Autism spectrum disorders (ASDs) encompass a broad range of developmental disorders that are marked by limitations in one of three behavioral/developmental domains: (1) social interaction; (2) language, communication, and imaginative play; and (3) range of interests and activities.[1] The ASDs range from phenotypically mild to severe and include autism, atypical autism, Asperger syndrome, Rett syndrome, and pervasive developmental disorders.[2] In 2007, the Centers for Disease Control and Prevention estimated that the prevalence of ASDs among children aged 8 years in the United States was ~1 in every 110 children.[3]

All of the disorders classified as ASDs are believed to be etiologically related,[4] possibly resulting from diverse changes in a common genetic pathway.[5] Results of twin and family studies have shown that the heritability of autism may be as high as 90%, making it one of the most heritable complex disorders.[6] In ~10% of patients, autism can be explained by monogenic disorders and known chromosomal abnormalities (e.g., the fragile X syndrome,[7,8] tuberous sclerosis,[9] Rett Syndrome,[10] and Potocki–Lupski syndrome[11]).[12]

Microdeletions and microduplications of chromosome 16p11.2 have been found at varying frequencies among individuals diagnosed with an ASD. The largest single study to date found that ~1% of autistic individuals carried either the 16p11.2 microdeletion or microduplication, as opposed to ~0.1% of individuals with a psychiatric or language disorder and ~0.01% of a population-based cohort which was not screened for psychiatric or language disorders.[13] The low prevalence of this genetic exposure, especially among controls, makes precisely estimating the magnitude of the association between 16p11.2 copy number variants (CNVs) and autism difficult. A meta-analysis of individuals with autism and/or developmental delay, combining patient-level data from three studies, estimated that 16p11.2 microdeletion is associated with a 38.7-fold increase in the odds of autism/developmental delay (OR, 38.7; 95% CI, 13.4–111.8). In addition, 16p11.2 microduplication was associated with a 20.7-fold increase in the odds of autism/developmental delay (OR, 20.7; 95% CI, 6.9–61.7).[14] The risk of autism/developmental delay was not significantly greater among patients with a 16p11.2 microdeletion when compared with those with the microduplication (OR, 2.04; 95% CI, 0.493–8.45, author's calculations).

Although such results firmly establish a role for the 16p11.2 microdeletion and microduplication in the etiology of autism and developmental delay, the extent to which these CNVs account for the total burden of ASDs has not been systematically investigated. Currently, the 1% of autism incidence typically attributed to CNVs at 16p11.2 is derived from a 2008 publication by Weiss et al.[13] This estimate makes CNVs at 16p11.2 at least as frequent as the most common known cytogenetic cause of autism (maternal 15q11–q13 duplication).[15]

Multiple studies of autistic subjects have investigated copy number variation in this genomic region. Some have found an equivalent proportion of 16p11.2 CNVs in those with ASDs as that reported by Weiss et al.,[16–18] whereas others have not.[19–22] Many studies that implicate these mutations in autism pathogenesis have used overlapping samples from collaborative biobanks, such as The Autism Genetic Resource Exchange.[5,13,14,17,18,20,22,23] As such, the autism-associated CNVs they report are frequently the second, third, or even fourth report of the same variants in the same individuals, a fact which frequently goes unmentioned. For instance, the ~0.7% frequency with which 16p11.2 CNVs were observed in individuals with an ASD by Kumar et al. has been repeatedly treated as a replication of the 1% frequency reported by Weiss et al., despite the fact that all probands with 16p11.2 CNVs in the study by Kumar et al. are also included in

From the Center for Perinatal Pediatric and Environmental Epidemiology, Yale School of Public Health, New Haven, Connecticut.

Kyle M. Walsh, CPPEE, One Church Street, 6th Floor, New Haven, CT 06510. E-mail: kyle.walsh@yale.edu.

Disclosure: The authors declare no conflict of interest.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.geneticsinmedicine.org).

Submitted for publication August 19, 2010.

Accepted for publication November 18, 2010.

Published online ahead of print February 1, 2011.

DOI: 10.1097/GIM.0b013e3182076c0c

Copyright © American College of Medical Genetics. Unauthorized reproduction of this article is prohibited.

the study by Weiss et al.[13,17] It is also likely that frequency estimates from Weiss et al. are inflated, as overestimation of allele frequency in a primary screen is a recognized phenomenon, especially in the investigation of rare variants.[24]

Despite large sample sizes in several individual studies, prevalence calculations for these particular CNVs are based on very few events, and confidence intervals are not reported. Furthermore, the study by Weiss et al. includes affected siblings with 16p11.2 CNVs in its prevalence calculations, despite sibling genotypes not being independent and their sample being enriched for multiplex families. Although probands from multiplex families have been shown to have fewer de novo CNVs than probands from simplex families,[16,18] little effort has been made to assess if the prevalence of 16p11.2 CNVs differs substantially between these family types.

A systematic literature review was conducted to identify studies that assayed copy number variation at chr16p11.2 among individuals with idiopathic ASDs using either a genome-wide scan or a targeted assay. A meta-analysis of the prevalence of 16p11.2 microdeletion/microduplication among autistic subjects was conducted to determine the proportion of ASDs attributable to this highly penetrant genetic variant. A planned subgroup analysis was also performed to test the hypothesis that these CNVs are more frequently found in cases of sporadic autism (simplex cases) than in cases of familial autism (multiplex cases). Methodological issues in the literature reporting on this association were addressed.

MATERIALS AND METHODS

Identification of eligible studies

Case-control studies and case-series reports that provided individual-level data were eligible for inclusion in the analysis if the total number of autistic individuals screened for CNVs at 16p11.2 was reported. Case reports were not eligible for inclusion. Genomewide studies of autistic patient populations were eligible if the study investigated copy number variation using either array comparative genomic hybridization or single nucleotide polymorphism genotyping platforms with sufficient resolution to detect variants $<$1Mb. Targeted studies of the 16p11.2 region were included if the CNVs were assayed by quantitative real-time polymerase chain reaction (PCR), multiplex ligation-dependent probe amplification, multiplex amplifiable probe hybridization, quantitative multiplex PCR of short fluorescent fragments, or metaphase fluorescent in situ hybridization.

A Medline search was performed on October 10, 2010, to identify all articles whose titles or abstracts contain the word "autism" or "autistic," and at least one of the following phrases: "structural variant," "genomic variant," "copy number," "deletion," "duplication," "microdeletion," or "microduplication." In addition, the bibliographies of selected articles were scanned to identify pertinent publications which the electronic search may have missed. Results were not filtered by language.

Article abstracts were scanned and eliminated from further analysis for the following reasons: ineligible phenotype, targeted analysis of incorrect genomic region, case report, review article, and basic science/animal model. Full-text versions of articles that passed the preliminary scan were downloaded and read in full. Articles were eliminated for the previously listed criteria in addition to the following: sample size not given, insufficient CNV detection resolution (fewer than 3 probes in 16p11.2 region), study prescreened and eliminated samples with 16p11.2 CNVs, and all case samples included in another study. Effort was made to include only unique patients; however, in studies where sample exclusivity could not be definitively ascertained, the study with the largest sample size was selected for analysis.

Phenotype specification

Individuals must have a physician-diagnosed ASD (autism, Asperger syndrome, or Pervasive Developmental Disorder-Not Otherwise Specified) according to Diagnostic and Statistical Manual of Mental Disorders, 4th edition or International Statistical Classification of Diseases and Related Health Problems 9/10 edition criteria or have received a diagnosis using a validated autism rating scale, including the Autism Diagnostic Interview-Revised, the Autism Diagnostic Observation Schedule, or the Childhood Autism Rating Scale. Participants were included irrespective of gender and ethnicity. When possible, individuals harboring gross chromosomal abnormalities or mutations in genes known to cause syndromic forms of autism (*FMR-1*, *MECP2*, and *TSC1*/*TSC2*) were excluded from analysis.

Prevalence calculations

Microdeletions and microduplications of chromosome 16p11.2 that span at least 65% of the recognized autism-associated region (chr16: 29,500,000–30,100,000) from the "Autism Chromosome Rearrangement Database" were defined as events.[16] CNVs that encompass this region but extend more than 1 Mb beyond these boundaries (either toward the centromere or telomere) were not included because such CNVs are associated with other known syndromes.[25–27]

Although not part of the primary analysis, de novo or maternally inherited duplications of chromosome 15q11–13 were recorded, as were all reported duplications of the chromosome 22q11.21 region. These duplications are well-described autism-associated variants, whose reciprocal deletions are associated with other neurological disorders (the Angelman/Prader-Willi syndromes and the DiGeorge/velocardiofacial syndromes, respectively).[15] We have abstracted data on these CNVs from six of seven studies included in the meta-analysis of CNVs at 16p11.2.

Using a fixed-effects model, a cumulative meta-analysis was conducted to estimate the prevalence of 16p11.2 microdeletion, 16p11.2 microduplication, and the combined prevalence of all pathogenic CNVs at this locus, among subjects with an ASD. 95% confidence intervals for these prevalence estimates were calculated. Study heterogeneity was assessed using the I^2 statistic.[28] Because different genetic mechanisms may be observed in sporadic cases compared with familial cases, a stratified analysis was conducted to determine if the prevalence of all 16p11.2 CNVs varied between these subgroups. A fixed-effects model was used to combine studies within each subgroup, and the study-to-study variance was computed independently within each subgroup. Heterogeneity between the two subgroups was assessed using the Q-value, which was compared with a χ^2 distribution with 1 degree of freedom using a significance threshold of $\alpha = 0.10$. All statistics were calculated in Comprehensive Meta-analysis, version 2 (Biostat Inc, Englewood, NJ).

Studies that did not indicate whether probands were from simplex or multiplex families were excluded from this subgroup analysis but appear in the cumulative meta-analyses. In multiplex families, only one individual from each family was included in the prevalence estimates. The included individual was the one defined by the authors as the proband. If probands were not identified, the eldest affected child in each family was assigned as proband.

© 2011 Lippincott Williams & Wilkins

Copyright © American College of Medical Genetics. Unauthorized reproduction of this article is prohibited.

Assessment of risk of bias in included studies

Included studies were determined to be free of substantial biases in case ascertainment. To eliminate bias as a result of low sensitivity, the genotyping platform used in a genomewide study was required to have at least 3 probes in the 16p11.2 region. To eliminate bias to low specificity, CNVs at 16p11.2 identified in a genomewide scan must have been confirmed using a second CNV-detection methodology (e.g., fluorescent in situ hybridization, quantitative real-time PCR, and multiplex ligation-dependent probe amplification) to avoid false positives.

To assess the presence of reporting bias, a funnel plot graphing study precision against the logit event rate was created and Egger's regression asymmetry test was performed.[29] By abstracting data on CNVs at 16p11.2 from genomewide scans, which are not hypothesis-driven, reporting bias should be minimized.

A sensitivity analysis was conducted by performing the meta-analysis with each study systematically omitted, one at a time with replacement, from the cumulative analysis to identify studies that had excessive influence on the summary measure.

RESULTS

The Medline search returned 503 results, of which 456 were excluded based on information available in the abstract. Scanning the citations of the 47 remaining articles identified an additional 4 articles which had not been returned by the Medline search and which could not be eliminated based solely on abstracts. Full-text versions of these 51 articles were obtained and scanned for relevance. After eliminating ineligible publications, a total of seven studies remained, all of which were retrieved as part of the Medline search (Fig. 1).[13,16,18–21,30] For one study, corresponding authors provided unpublished genomic coordinates for CNVs detected in the 16p11.2 region.[21]

The study by Glessner et al.[20] did not report excluding patients with known syndromic forms of autism. As such, the prevalence of 16p11.2 CNVs in this cohort likely underestimates the true prevalence of these CNVs among patients with idiopathic ASDs. 17% of ASD cases reported by Guilmatre et al.[19] were not diagnosed using a standardized scale and instead received autism diagnoses at units specializing in the diagnosis and evaluation of ASDs. All cases reported in this study were included in the meta-analyses.

Characteristics of studies included in the meta-analysis can be seen in Table 1 and a breakdown of all individuals included in the meta-analysis, by study, can be seen in Table 2. Some studies used unique discovery cohorts, but replication cohorts which overlapped with samples used in different studies. Other studies used overlapping samples in their discovery cohorts, but unique replication cohorts. After excluding overlapping samples, and siblings of probands, a total of 3613 individuals with ASDs were included in the meta-analysis.

The estimated prevalence of 16p11.2 microdeletion/microduplication among all idiopathic ASD probands was 0.76% (95% CI, 0.51–1.12%; Fig. 2). Results of a sensitivity analysis, conducted by performing the cumulative meta-analysis with each study systematically omitted, one at a time with replacement, did not indicate that any one study was exerting undue influence on the summary measure. In addition, no substantial study heterogeneity was detected in the cumulative meta-analysis ($I^2 = 0$).

When stratifying by CNV-type, the estimated prevalence of 16p11.2 microdeletions among all idiopathic ASD probands was 0.50% (95% CI, 0.31–0.82%). The estimated prevalence of 16p11.2 microduplications among all idiopathic ASD probands was 0.28% (95% CI, 0.14–0.56%; Fig. 3). Although these results indicate that the deletion may be a more penetrant variant than the duplication, these prevalence estimates did not differ significantly ($Q = 1.83$, $P = 0.176$).

Five of seven studies reported whether individuals were from simplex or multiplex families and were included in a subgroup analysis. The studies by Glessner et al.[20] and Shen et al.[30] did not provide these data and were excluded from subgroup analysis. Based on data from four studies, the prevalence of 16p11.2 microdeletion/microduplication among idiopathic ASD probands from multiplex families is 0.76% (95% CI, 0.36–1.58%; $I^2 = 0$). Based on data from these four studies, plus one additional study, the prevalence of 16p11.2 microdeletion/microduplication among idiopathic ASD probands from simplex

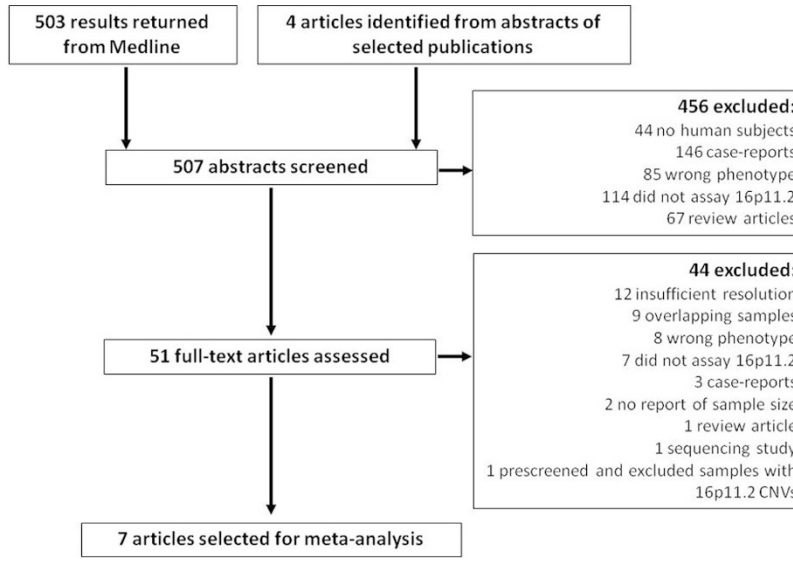

Fig. 1. Flow of published articles through the different phases of the review.

Copyright © American College of Medical Genetics. Unauthorized reproduction of this article is prohibited.

Table 1 Characteristics of included studies

Study	Source population	Percent male	ASD diagnostic criteria	Exclusion criteria	Primary CNV detection method	CNV confirmation method
Guilmatre et al.[19]	European[a]	80.5	ADI-R, ADOS, or CARS[b]	Abnormal karyotype, *FMR1* mutations[c]	QMPSF	Array CGH
Marshall et al.[16]	Canadian	—	ADI-R and ADOS	Abnormal karyotype, *FMR1* mutations[d]	SNP array	QMPSF or qPCR
Sebat et al.[17]	American	83.6	ADI-R and ADOS	Abnormal karyotype, *FMR1* mutations	ROMA	FISH
Weiss et al.[13]	American/Icelandic[e]	—	ADI-R and ICD-10	Abnormal karyotype, *FMR1* mutations	SNP array	MLPA
Van der Zwaag et al.[21]	Dutch	—	ADI-R and ADOS	Abnormal karyotype, *FMR1* or *TSC1/2* mutations	SNP array	qPCR
Glessner et al.[20][f]	Caucasian Americans	81.8	ADI-R	Abnormal karyotype, *FMR1* mutations	SNP array	qPCR or MLPA
Shen et al.[30]	American	80.9	DSM-IV	Abnormal karyotype, *FMR1* mutations	Array CGH[g]	Secondary array

[a]Includes 231 French patients and 29 Italian patients.
[b]17% of ASD cases not diagnosed using a standardized scale.
[c]66% of ASD cases screened.
[d]75% of ASD cases screened.
[e]Includes 751 American patients and 299 Icelandic patients.
[f]Data in the table refer only to the primary cohort, and not the AGRE replication cohort.
[g]12.7% assayed using either the Affy 500K of Affy 5.0 SNP array.
ADI-R, Autism Diagnostic Interview–Revised; ADOS, the Autism Diagnostic Observation Schedule; CARS, the Childhood Autism Rating Scale; ICD-10, International Statistical Classification of Diseases and Related Health Problems 10th Revision; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders 4th Edition; *FMR1*, fragile X mental retardation 1; *TSC1/2,* Tuberous sclerosis 1 and 2; QMPSF, Quantitative Multiplex Polymerase Chain Reaction of Short Fluorescent Fragments; array-CGH, array-comparative genomic hybridization; qPCR, quantitative polymerase chain reaction; ROMA, representational oligonucleotide microarray analysis; FISH, fluorescent in situ hybridization; MLPA, multiplex ligation-dependent probe amplification; SNP, single nucleotide polymorphism.

families is 1.13% (95% CI, 0.54–2.35%; $I^2 = 0$; Fig. 4). As expected, the prevalence of CNVs at 16p11.2 was higher among simplex cases than among multiplex cases when first-degree relatives were excluded from analysis. Despite this difference in prevalence between subgroups, significant heterogeneity was not detected ($Q = 0.563$, $P = 0.453$), supporting the calculation of a single summary prevalence estimate.

The presence of reporting bias was assessed by visually inspecting a funnel plot, which graphed study precision against the logit event rate (Fig. 5). The funnel plot showed no evidence of reporting bias, although there were no studies in the lower right corner. This region corresponds to studies with low precision (small sample size) and a large logit event rate. Absence of studies in this region of the funnel plot is expected, as small studies are unlikely to detect very rare variants. The Egger et al.[29] regression asymmetry test did not suggest that publication bias was present (Intercept = −0.188, $P = 0.565$). However, because of the limited number of studies in the analysis, this result should be treated with caution.

DISCUSSION

The prevailing hypothesis for the genetic etiology of autism has largely been the same as that for other common diseases and is widely referred to as the "common gene/common disease" hypothesis. The theory is that common diseases result from the additive or multiplicative effects of genetic and environmental factors. Under this paradigm, common genetic variants confer only a small increased risk to a given individual, but because of the high frequency with which these variants are found, each has a large attributable risk among the population. This hypothesis is readily tested using genome-wide association studies, and such studies have had some successes in unraveling autism biology.[31–33]

An alternative to the "common gene/common disease" hypothesis is that ASDs are caused not only by common variants of small effect but also by rare highly penetrant variants such as chromosomal deletions and duplications.[34] With the development of high-resolution microarrays, it is now possible to scan the genomes of autistic individuals for pathogenic CNVs that escape detection by previous-generation cytogenetic methodologies.

A substantial proportion of idiopathic autism may be attributable to CNVs.[34] Two recent studies detected de novo CNVs in 7–10% of autistic cases from simplex families, 2–3% of cases from multiplex families, and in 1% of controls.[16,18] These results not only implicate CNVs in the etiology of autism but also indicate that different genetic mechanisms may underlie sporadic, versus familial, autism.

Recent studies of patients with ASDs have identified several recurrent, but rare, CNVs that were absent or markedly less common in controls.[34] Some of these CNVs are located in chromosomal regions associated with other neurologic disorders, such as the Prader-Willi and Angelman syndromes (Chr15q11–13) and the DiGeorge/velocardiofacial syndrome (22q11.21).[15] Six of the seven studies included in the meta-analysis contained data that permitted calculation of the prevalence of both Chr15q11–13 duplications (maternally inherited

 © 2011 Lippincott Williams & Wilkins

Copyright © American College of Medical Genetics. Unauthorized reproduction of this article is prohibited.

Table 2 Breakdown of all individuals included in the meta-analysis, by study and family-type

Study	ASD cohorts	Sample size	Unique idiopathic ASD probands[a]	16p11.2 deletions	16p11.2 duplications	Frequency
Guilmatre et al.[19]	Simplex	25	25	0	0	0/25
	Multiplex	235	235	0	0	0/235
	Total	260	260	0	0	0/260
Marshall et al.[16]	Simplex	238	214	1	2	3/214
	Multiplex	189	181	1	0	1/181
	Total	427	395[b]	2	2	4/395
Sebat et al.[17]	Simplex	118	118	1	0	1/118
	Multiplex	77	0[c]	—	—	—
	Total	195	118	1	0	1/118
Weiss et al.[13]	Simplex	299	299	3	0	3/299
	Multiplex	1441	751[d]	3[e]	3	6/751
	Total	1740	1050	6	3	9/1050
Van der Zwaag et al.[21]	Simplex	53	53	0	0	0/53
	Multiplex	52	52	0	0	0/52
	Total	105	105	0	0	0/105[f]
Glessner et al.[20g]	Simplex	531	464	—	—	—
	Multiplex	1664	395	—	—	—
	Total	2195	859[h]	3	2	5/859
Shen et al.[30i]	Simplex	—	—	—	—	—
	Multiplex	—	—	—	—	—
	Total	848	826[j]	4	1	5/826

[a]These are the individuals included in the meta-analysis after eliminating: overlapping samples, affected relatives of the proband, samples with gross chromosomal abnormalities, and when possible, samples with mutations in genes known to cause syndromic forms of autism (*FMR-1* and *TSC1*/*TSC2*).
[b]32 samples were excluded due to the presence of an abnormal karyotype.
[c]AGRE samples which overlap with those in Weiss et al.[13]
[d]751 unique probands existed after eliminating affected kindreds.
[e]After assigning 751 unique probands, 3 were determined to have 16p11.2 deletions. An additional multiplex ASD family was found to have this deletion, but does not appear here because it was detected in the proband's affected sibling, but not in the proband.
[f]Data on 16p11.2 deletions and duplications were obtained through correspondence with the author (B. van der Zwaag). No 16p11.2 CNVs detected in this cohort met inclusion criteria due to insufficient overlap with critical region.
[g]Authors do not indicate whether the identified CNVs were detected in multiplex or simplex families.
[h]A total of 859 individuals were included in the meta-analysis. The excluded individuals (from Glessner et al's replication cohort[20]) were part of the AGRE sample and overlap with samples in Weiss et al.[13]
[i]Authors do not indicate the number of multiplex or simplex families.
[j]18 samples were excluded due to the presence of an abnormal karyotype and 4 were excluded due to the presence of *FMR-1* mutations or premutations.

and de novo only) and 22q11.21 duplications among patients with idiopathic ASDs. Maternally inherited or de novo duplication of 15q11–13 was observed in 0.70% of patients (95% CI, 0.45–1.10%), whereas duplication of 22q11.21 was observed in 0.60% of patients (95% CI, 0.34–1.06%; Fig., Supplemental Digital Content 1, http://links.lww.com/GIM/A142). These results are not from a meta-analysis specifically designed to assess the prevalence of CNVs in these regions. However, within the cohorts under study, microdeletion/microduplication of 16p11.2 was a more common cause of ASDs than duplication of these other regions.

The prevalence of 16p11.2 microdeletion/microduplication among ASD probands is typically given as 1%, despite this estimate being largely based on results from two studies with overlapping samples.[13,17] Because of the limited samples sizes, use of overlapping samples, and several other issues summarized in Table 3, a systematic review and meta-analysis of 3613 ASD patients was performed. The meta-analysis determined that the prevalence of 16p11.2 microdeletion/microduplication among ASD probands is 0.76% (95% CI, 0.51–1.12%), and that microdeletions are a more common cause of ASDs than the reciprocal microduplication (0.50% vs. 28%, respectively). Although the subgroup analysis suggests that the prevalence of 16p11.2 CNVs may be higher in cases of sporadic autism than in cases of familial autism, this difference did not reach statistical significance ($P = 0.453$).

Although seven studies is a moderate-sized sample for a meta-analysis, it is important to use strictly defined phenotypes to reach valid conclusions. Although studies have identified 16p11.2 CNVs in individuals with multiple cognitive abnormal-

Copyright © American College of Medical Genetics. Unauthorized reproduction of this article is prohibited.

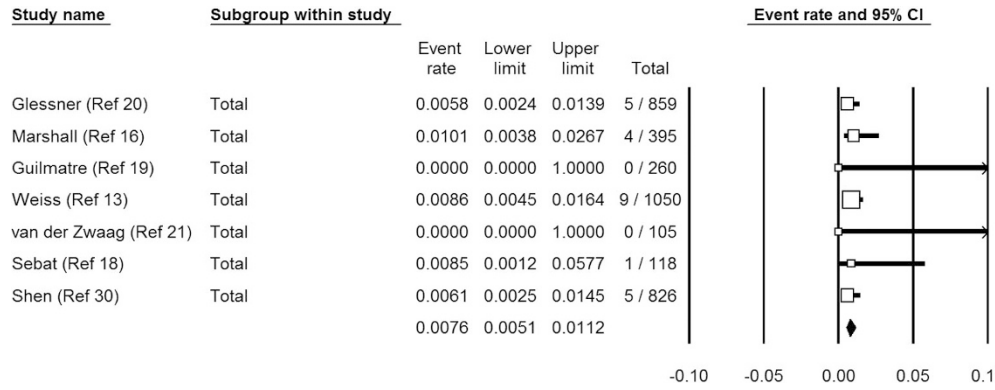

Fig. 2. Forest plot of the overall prevalence of 16p11.2 microdeletion/microduplication among all ASD probands, irrespective of family-type (i.e., simplex/sporadic cases versus multiplex/familial cases) or CNV-type (deletion versus duplication). The area of study symbols (white squares) is proportional to study weight.

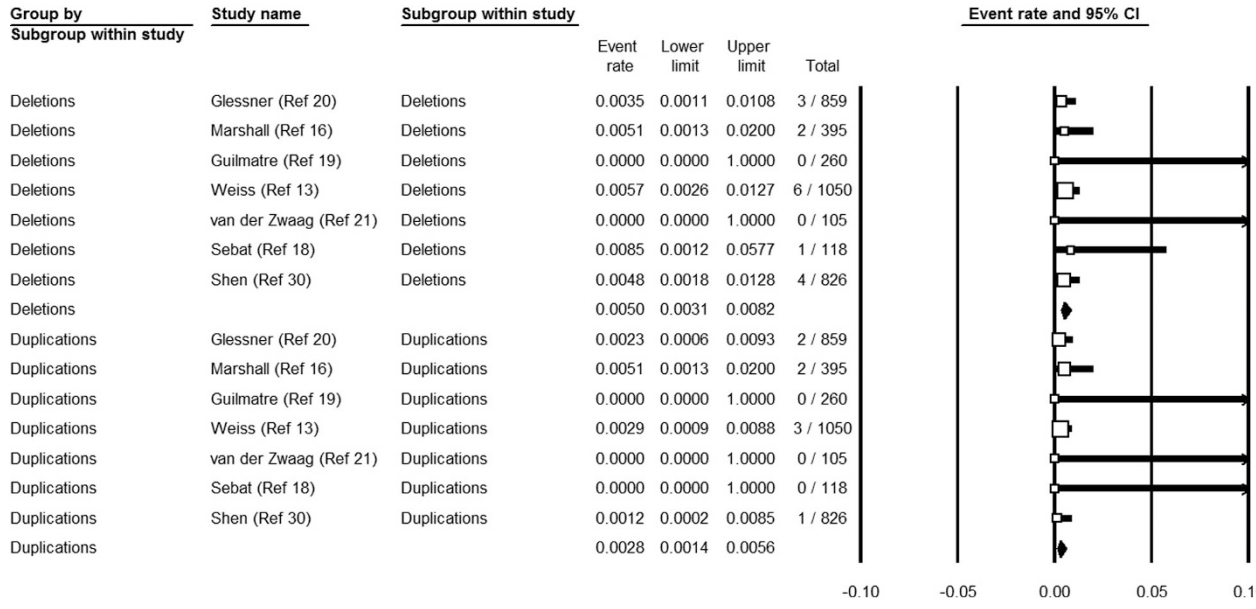

Fig. 3. Forest plots of the prevalence of 16p11.2 microdeletions and 16p11.2 microduplications (stratified by CNV-type) among all ASD probands. The area of study symbols (white squares) is proportional to study weight.

ities, developmental delay, and mental retardation, these studies were not included because the meta-analysis was specifically designed to investigate the proportion of idiopathic ASDs attributable to these structural variants.[14,35–38] A recent study of patients referred to a clinical genetics testing center with an ASD as the primary indication found that 7 of 580 individuals carried a 16p11.2 CNV (5 deletions and 2 duplications).[39] Despite an ASD being the indication for referral, it was unknown if any diagnostic rating scale was used to obtain this diagnosis. Furthermore, it was unknown what criteria were used to make the diagnosis (e.g., DSM-IV). Therefore, this study was excluded from the meta-analysis. It is interesting to note that the total prevalence of 16p11.2 CNVs in this sample (1.21%) is slightly higher than the upper 95% CI returned from the meta-analysis (1.12%), possibly due to phenotypic misspecification.

To properly estimate the proportion of multiplex ASD subjects who carry the 16p11.2 microdeletion/microduplication, it is vital that only one member of each family be included. Failure to ensure this will cause a unit of analysis error and can produce inflated prevalence estimates. Choosing which member of a multiplex family to include should be done in an unbiased way, without prior knowledge of the sibling genotypes. Although this may seem a minor point, four independent families in the Weiss et al.[13] study had multiple affected children, at least one of whom carried a 16p11.2 CNV and at least one of whom did not. Thus, as few as zero and as many as four additional 16p11.2 CNVs could enter the meta-analysis based solely on which children are selected as the probands. Based on our predetermined criteria, three of these four individuals were included.

The ratio of multiplex cases to simplex cases in this meta-analysis is artificially elevated because families with multiple affected children have been preferentially recruited for genetic studies. Because of this, the ratio of familial to sporadic ASD probands in the meta-analysis does not accurately reflect the true ratio among all ASD subjects. As the subgroup analysis did not detect hetero-

 © 2011 Lippincott Williams & Wilkins

Copyright © American College of Medical Genetics. Unauthorized reproduction of this article is prohibited.

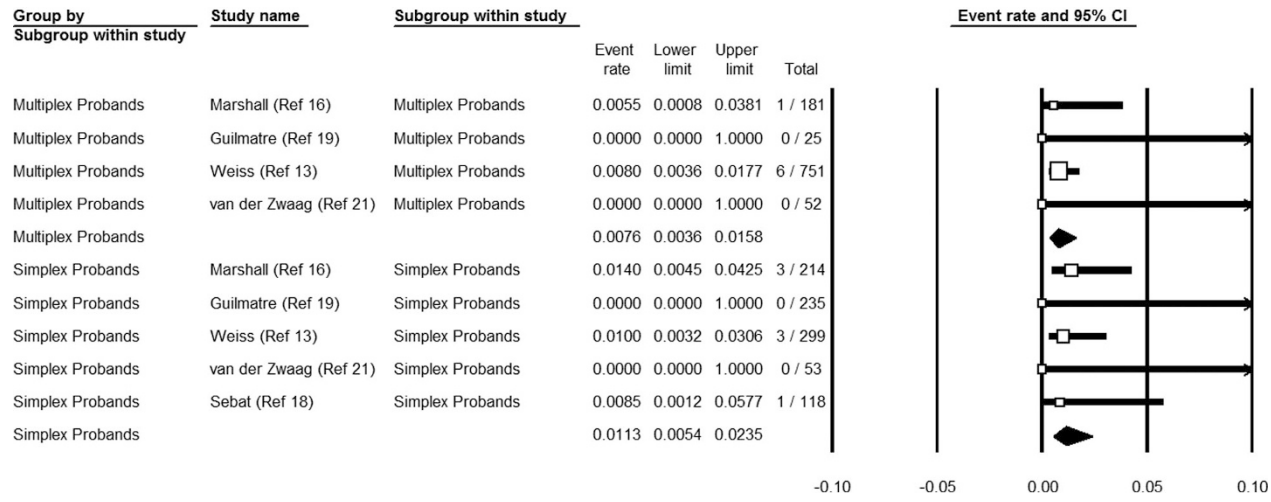

Fig. 4. Forest plot of the overall prevalence of 16p11.2 microdeletion/microduplication among idiopathic ASD probands, stratified by family-type (i.e., simplex/sporadic cases versus multiplex/familial cases). The area of study symbols (white squares) is proportional to study weight.

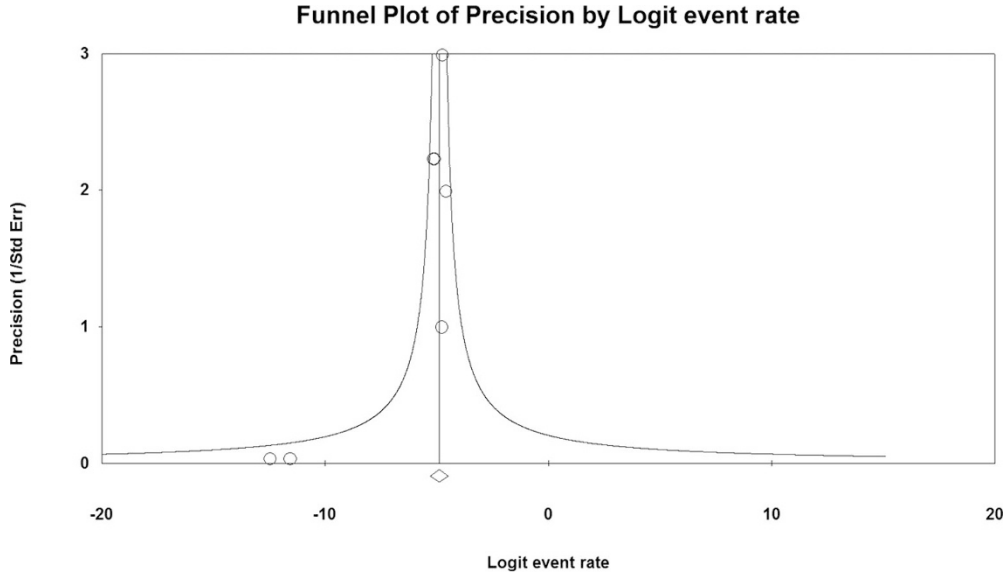

Fig. 5. A funnel plot was constructed by graphing study precision against the logit event rate to visually assess the presence or absence of publication bias.

geneity between these groups, the prevalence estimated here should accurately reflect that among all ASD probands.

If true heterogeneity in the prevalence of these CNVs does exist across family type, it may not be detected here because, in this subgroup analysis, duplications and deletions were analyzed jointly. The lower penetrance microduplication is likely more common in multiplex families than the microdeletion, as it has been identified in carrier parents who can pass it to multiple affected children.[13] Of note, 16p11.2 deletions were also detected in multiplex families. In one of these families, only one affected child carried the deletion, highlighting the genetic heterogeneity of ASDs.[16] In a second family, two children carried a de novo deletion, indicating that parental germline mosaicism may also account for the similar prevalence in multiplex and simplex probands.[13]

A recent systematic review concluded that chromosomal microarray should be the first-tier cytogenetic diagnostic test for patients with developmental delay, multiple cognitive abnormalities, or ASDs.[40] However, only 2 of the 33 publications included in that review analyzed patients with ASDs. If 16p11.2 microdeletion/microduplication is in fact the most common cytogenetic cause of autism, its prevalence among individuals with ASDs will greatly influence the diagnostic yield of routine microarray analysis performed on these patients. It is vital to know what proportion of autism incidence may be explained by these mutations when considering the utility of screening for microdeletions and microduplications among autistic patient populations, especially in clinical settings.

If 0.76% of individuals with an idiopathic ASD have either the 16p11.2 microdeletion or the reciprocal microduplication,

Copyright © American College of Medical Genetics. Unauthorized reproduction of this article is prohibited.

Table 3 Weaknesses in previous prevalence estimates of 16p11.2 microdeletion/microduplication among individuals with an idiopathic ASD

Overestimation of allele frequency in a primary screen is especially common in the investigation of rare variants.[24]
Prevalence calculations based on a small number of events will be highly variable and have wide confidence intervals.
Sibling genotypes are not independent. Therefore, only one member of each family should be included when using a family-based sample to estimate the frequency of a genetic variant.
The frequency of a high-penetrance variant may differ in probands with sporadic, versus familial, disease.
To avoid publication bias, data from other studies should be considered, especially studies which do not report an increased 16p11.2 CNV frequency among cases but do have the ability to detect these variants.
Studies used to assess 16p11.2 CNV frequency often use overlapping samples.

the number needed to test to identify one such variant is 132 patients (95% CI, 89–198). Such prevalence information, especially as it pertains to diagnostic yield in genetic testing, should prove useful to clinicians considering chromosomal microarray analysis in subjects with ASDs.

ACKNOWLEDGMENTS

The authors thank B. van der Zwaag and Maretha V. de Jonge of the Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht for providing unpublished genomic coordinates for copy number variants reported in the study by van der Zwaag et al.[21]

REFERENCES

1. Muhle R, Trentacoste SV, Rapin I. The genetics of autism. *Pediatrics* 2004;113:e472–e486.
2. Sykes NH, Lamb JA. Autism: the quest for the genes. *Expert Rev Mol Med* 2007;9:1–15.
3. Autism and Developmental Disabilities Monitoring Network Surveillance Year 2006 Principal Investigators; Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders—Autism and Developmental Disabilities Monitoring Network, United States, 2006. *MMWR Surveill Summ* 2009;58:1–20.
4. Reichenberg A, Gross R, Weiser M, et al. Advancing paternal age and autism. *Arch Gen Psychiatry* 2006;63:1026–1032.
5. Pinto D, Pagnamenta AT, Klei L, et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 2010;466:368–372.
6. Freitag CM. The genetics of autistic disorders and its clinical relevance: a review of the literature. *Mol Psychiatry* 2007;12:2–22.
7. Verkerk AJ, Pieretti M, Sutcliffe JS, et al. Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 1991;65:905–914.
8. Yu S, Pritchard M, Kremer E, et al. Fragile X genotype characterized by an unstable region of DNA. *Science* 1991;252:1179–1181.
9. European Chromosome 16 Tuberous Sclerosis Consortium. Identification and characterization of the tuberous sclerosis gene on chromosome 16. *Cell* 1993;75:1305–1315.
10. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999;23:185–188.
11. Potocki L, Bi W, Treadwell-Deering D, et al. Characterization of potocki-lupski syndrome (dup(17)(p11.2p11.2)) and delineation of a dosage-sensitive critical interval that can convey an autism phenotype. *Am J Hum Genet* 2007;80:633–649.
12. Abrahams BS, Geschwind DH. Advances in autism genetics: on the threshold of a new neurobiology. *Nat Rev Genet* 2008;9:341–355.
13. Weiss LA, Shen Y, Korn JM, et al. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 2008;358:667–675.
14. McCarthy SE, Makarov V, Kirov G, et al. Microduplications of 16p11.2 are associated with schizophrenia. *Nat Genet* 2009;41:1223–1227.
15. Cook EH Jr, Scherer SW. Copy-number variations associated with neuropsychiatric conditions. *Nature* 2008;455:919–923.
16. Marshall CR, Noor A, Vincent JB, et al. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008;82:477–488.
17. Kumar RA, KaraMohamed S, Sudi J, et al. Recurrent 16p11.2 microdeletions in autism. *Hum Mol Genet* 2008;17:628–638.
18. Sebat J, Lakshmi B, Malhotra D, et al. Strong association of de novo copy number mutations with autism. *Science* 2007;316:445–449.
19. Guilmatre A, Dubourg C, Mosca AL, et al. Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic pathways in schizophrenia, autism, and mental retardation. *Arch Gen Psychiatry* 2009; 66:947–956.
20. Glessner JT, Wang K, Cai G, et al. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* 2009;459:569–573.
21. van der Zwaag B, Franke L, Poot M, et al. Gene-network analysis identifies susceptibility genes related to glycobiology in autism. *PLoS One* 2009;4: e5324.
22. Bucan M, Abrahams BS, Wang K, et al. Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS Genet* 2009;5:e1000536.
23. Christian SL, Brune CW, Sudi J, et al. Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder. *Biol Psychiatry* 2008; 63:1111–1117.
24. Hirschhorn JN, Altshuler D. Once and again-issues surrounding replication in genetic association studies. *J Clin Endocrinol Metab* 2002;87:4438–4441.
25. Ballif BC, Hornor SA, Jenkins E, et al. Discovery of a previously unrecognized microdeletion syndrome of 16p11.2-p12.2. *Nat Genet* 2007;39:1071–1073.
26. Battaglia A, Novelli A, Bernardini L, Igliozzi R, Parrini B. Further characterization of the new microdeletion syndrome of 16p11.2-p12.2. *Am J Med Genet A* 2009;149A:1200–1204.
27. Bochukova EG, Huang N, Keogh J, et al. Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature* 2010;463:666–670.
28. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–560.
29. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629–634.
30. Shen Y, Dies KA, Holm IA, et al. Clinical genetic testing for patients with autism spectrum disorders. *Pediatrics* 2010;125:e727–e735.
31. Weiss LA, Arking DE, Gene Discovery Project of Johns Hopkins, the Autism Consortium, Daly MJ, Chakravarti A. A genome-wide linkage and association scan reveals novel loci for autism. *Nature* 2009;461:802–808.
32. Ma D, Salyakina D, Jaworski JM, et al. A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *Ann Hum Genet* 2009;73(Pt 3):263–273.
33. Autism Genome Project Consortium, Szatmari P, Paterson AD, et al. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* 2007;39:319–328.
34. Kusenda M, Sebat J. The role of rare structural variants in the genetics of autism spectrum disorders. *Cytogenet Genome Res* 2008;123:36–43.
35. Bijlsma EK, Gijsbers AC, Schuurs-Hoeijmakers JH, et al. Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals. *Eur J Med Genet* 2009;52:77–87.
36. Shinawi M, Liu P, Kang SH, et al. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioral problems, dysmorphism, epilepsy, and abnormal head size. *J Med Genet* 2010;47:332–341.
37. Mefford HC, Cooper GM, Zerr T, et al. A method for rapid, targeted CNV genotyping identifies rare variants associated with neurocognitive disease. *Genome Res* 2009;19:1579–1585.
38. Baldwin EL, Lee JY, Blake DM, et al. Enhanced detection of clinically relevant genomic imbalances using a targeted plus whole genome oligonucleotide microarray. *Genet Med* 2008;10:415–429.
39. Rosenfeld JA, Ballif BC, Torchia BS, et al. Copy number variations associated with autism spectrum disorders contribute to a spectrum of neurodevelopmental disorders. *Genet Med* 2010;12:694–702.
40. Miller DT, Adam MP, Aradhya S, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 2010;86: 749–764.

 © 2011 Lippincott Williams & Wilkins

Copyright © American College of Medical Genetics. Unauthorized reproduction of this article is prohibited.