RESEARCH ARTICLE

Patterns of Risk for Multiple Co-Occurring Medical Conditions Replicate Across Distinct Cohorts of Children with Autism Spectrum Disorder

Kimberly A. Aldinger, Christianne J. Lane, Jeremy Veenstra-VanderWeele, and Pat Levitt

Children with autism spectrum disorder (ASD) may present with multiple medical conditions in addition to ASD symptoms. This study investigated whether there are predictive patterns of medical conditions that co-occur with ASD, which could inform medical evaluation and treatment in ASD, as well as potentially identify etiologically meaningful subgroups. Medical history data were queried in the multiplex family Autism Genetic Resource Exchange (AGRE). Fourteen medical conditions were analyzed. Replication in the Simons Simplex Collection (SSC) was attempted using available medical condition data on gastrointestinal disturbances (GID), sleep problems, allergy and epilepsy. In the AGRE cohort, no discrete clusters emerged among 14 medical conditions. GID and seizures were enriched in unaffected family members, and together with sleep problems, were represented in both AGRE and SSC. Further analysis of these medical conditions identified predictive co-occurring patterns in both samples. For a child with ASD, the presence of GID predicts sleep problems and vice versa, with an approximately 2-fold odds ratio in each direction. These risk patterns were replicated in the SSC sample, and in addition, there was increased risk for seizures and sleep problems to co-occur with GID. In these cohorts, seizure alone was not predictive of the other conditions co-occurring, but behavioral impairments were more severe as the number of co-occurring medical symptoms increased. These findings indicate that interdisciplinary clinical care for children with ASD will benefit from evaluation for specific patterns of medical conditions in the affected child and their family members. *Autism Res* 2015, 8: 771–781. © 2015 The Authors Autism Research published by Wiley Periodicals, Inc. on behalf of International Society for Autism Research.

Keywords: gastrointestinal disturbances; sleep; seizure; medical symptoms

Introduction

Advancing interdisciplinary clinical care for children with autism spectrum disorders (ASD) poses significant challenges. Although defined through a categorical diagnosis, individuals can present with highly variable symptoms related to the core features of ASD: social communication dysfunction and restricted, repetitive behavior (American Psychiatric Association, 2013; Sacco et al., 2010). Moreover, certain medical conditions more commonly occur in children with an ASD diagnosis (Geschwind, 2009; Isaksen et al., 2012; Memari, Ziaee, Mirfazeli, & Kordi, 2013; Sacco et al., 2010). As a highly heritable, polygenic disorder, genetic studies have identified many ASD-associated risk genes and copy number variants (Berg, & Geschwind, 2012; Murdoch, & State, 2013; Scherer, & Dawson, 2011). However, these discoveries have yet to translate into predictable clinical subtypes except when intellectual disability is also present (Jeste, & Geschwind, 2014).

In an effort to reduce phenotypic heterogeneity in the context of genetic studies, specific patterns of core ASD symptoms, including social communication and restrictive and repetitive behaviors, have been used to attempt to identify biologically meaningful subgroups within ASD (Bradford et al., 2001; Campbell et al., 2009; Sacco et al., 2010; Silverman et al., 2002; Veatch, Veenstra-Vanderweele, Potter, Pericak-Vance, & Haines,

2014). This strategy has had mixed results, perhaps in part because of the underlying structures of diagnostic or screening instruments that were developed to identify people with ASD rather than to parse differences among individuals with ASD (Boomsma et al., 2008; Constantino et al., 2004; Constantino, 2002; Frazier, Youngstrom, Kubu, Sinclair, & Rezai, 2008; Lecavalier, 2006; Lord, Rutter, DiLavore, & Risi, 1999; Rutter, Le Couteur, & Lord, 2003; Tadevosyan-Leyfer et al., 2003). An alternative approach would be to evaluate potential subgroups within ASD using co-occurring medical conditions that may be more clearly defined than ASD itself. This approach is supported by the observation that clusters of some medical conditions are indicative of genetic syndromes, such as fragile X syndrome, Rett syndrome, 16p11.2 deletion and 22q deletion that markedly increase the risk of ASD (Jeste, & Geschwind, 2014; Kohane et al., 2012). However, these syndromes are rare in the ASD population, and the unique patterns of comorbidities ascribed to each syndrome may not relate to those medical conditions that most commonly occur in nonsyndromal ASD (Gaugler et al., 2014).

Numerous medical conditions may co-occur in children with ASD, making it challenging to fully ascertain medical conditions in large datasets (Buie et al., 2010; Campbell et al., 2009; Gorrindo et al., 2012; Hilton, Zhang, Whilte, Klohr, & Constantino, 2012; Kohane et al., 2012; Lajonchere, & Consortium, 2010; McElhanon, McCracken, Karpen, & Sharp, 2014; Sacco et al., 2010; Sikora, Johnson, Clemons, & Katz, 2012). The prevalence of medical disturbances in ASD is typically reported over a wide range, likely corresponding both to challenges in ascertaining these conditions as well as differences between ASD samples (Geschwind, 2009; Kohane et al., 2012). Examining the relationships among co-occurring medical conditions in children with ASD that have been ascertained in large samples may provide initial evidence for clustering of medical conditions in ASD, while offering an opportunity to identify patterns of co-occurrence within a child, thus motivating a broader medical assessment in future studies. Specific patterns of co-occurring medical conditions also may reflect a shared biological load among subgroups of children with ASD, which may also be related to ASD symptom severity. This has been addressed to a limited extent for gastrointestinal disturbances (GID; Gorrindo et al., 2012; Peters et al., 2013). More recently, medical conditions extracted from medical record data at Boston Children's Hospital identified three ASD subgroups characterized by seizures, multisensory dysfunction, and psychiatric disorders (Doshi-Velez, Ge, & Kohane, 2014); although the majority of the sample (91.6%) did not resolve into any of these three subgroups.

Here, we report the results of performing hierarchical clustering and multivariate analyses on two large samples of multiplex and simplex families, the Autism Genetic Resource Exchange (AGRE) and the Simons Simplex Collection (SSC), respectively. These two datasets include information on the presence of GID, seizure disorders, allergies, and sleep problems. Based upon previous literature showing a relationship between GID and sleep problems, as well as seizure disorders and co-occurring medical conditions, we hypothesized that symptoms in any one of these domains would increase the likelihood of symptoms in the others. Furthermore, we hypothesized that increasing co-occurring medical problems would be associated with increasing behavioral difficulties. The AGRE database includes information on a large variety of co-occurring medical conditions, and we hypothesized that a cluster analysis within this richer dataset would allow identification of subgroups defined by specific co-occurring conditions. Beyond the identification of biologically meaningful relationships between ASD and co-occurring medical conditions, these analyses may also reveal important clinical relationships that inform assessment and treatment of children with ASD.

Methods
Study Samples

ASD family participants were collected by the Autism Genetics Research Exchange (AGRE) consortium (Geschwind et al., 2001). The AGRE collection is a family based cohort that predominantly consists of families with more than one child with ASD (multiplex). The following datasets were downloaded with permission from the AGRE web site (http://www.research.agre.org): (1) pedigree file, (2) medical history file, (3) Autism Diagnostic Interview-Revised (ADI-R), (4) the parent report form of the Social Responsiveness Scale (SRS), (5) the Vineland, and (6) Stanford-Binet Intelligence Scale. The pedigree file contained information on family structure and autism diagnosis. The medical history files contained the child's medical history collected through parent report. The ADI-R is a semistructured interview used to assess autism symptoms in support of clinical diagnosis (Rutter et al., 2003). The SRS is a continuous rating scale completed by parents or teachers used to assess severity of social impairments within a population (Constantino, 2002). The Vineland Adaptive Behavior Scales II (VABS-II) is a parent report questionnaire of adaptive behavior across the lifespan (Sparrow, Cicchetti, & Balla, 2005). The VABS-II content is appropriate for individuals with a wide range of abilities, including those who are nonverbal or have intellectual disabilities, and includes a communication domain. The AGRE cohort that had detailed medical information included 728 individuals with ASD from 351 families (95.2% multiplex).

Table 1. Demographics and Medical Symptoms

		AGRE cohort					SSC Cohort	
		Multiplex cohort		Cluster Analysis sample				
N		728		627			2623	
Child Characteristics		M	SD	M	SD		M	SD
Age		9.2	5.1	9.3	5.2		9.0	3.6
Vineland Composite		52.6	0.2	53.8	0.2	***		
		N	%	N	%		N	%
Sex	Male	570	78.3%	500	79.7%	*	2275	86.7%
	Female	158	21.7%	127	20.3%		348	13.3%
European Race		601	82.6%	507	80.9%	**	2205	85.7%
Hispanic Ethnicity		108	14.8%	94	15.0%		14	0.6%
Medical Conditions								
Abnormal growth		75	10.7%	66	10.5%		—	—
Allergies		291	40.0%	284	45.3%	***	891	34.0%
Asthma		113	16.3%	108	17.2%	*	—	—
Coordination		291	40.0%	273	43.5%	***	—	—
Gait abnormalities		179	24.6%	160	25.5%		—	—
Gastrointestinal		309	42.4%	279	44.5%	**	1121	43.1%
Movement		94	12.9%	86	13.7%		—	—
Respiratory problems		80	11.4%	75	12.0%		—	—
Seizures		89	12.2%	70	11.2%	*	129	5.3%
Skin abnormalities		218	30.8%	206	32.9%	***	—	—
Sleep disorder		361	55.5%	345	55.0%		1896	72.5%
Stereotypies		539	77.4%	486	77.5%		—	—
Teeth abnormalities		103	14.6%	95	15.2%		—	—
Vision problems		103	14.9%	96	15.3%		—	—

Note. Specific medial abnormalities are dichotomized by presence of any abnormality (hypersensitivity or hyposensitivity).
Asterisks refer to comparison of ASD sample included in cluster analysis with those not included: * $P < 0.05$; **$P < 0.01$; ***$P < 0.001$.
Percentages reported exclude unknown diagnoses.

The SSC is a family-based cohort that consists of families having one child with ASD (simplex). The following datasets were downloaded with permission from the SFARI Base web site (http://sfari.org/resources/sfari-base) for public cohort 13: (1) demographics, (2) medical history form, (3) parent report form of the SRS (4) VABS-II, and (5) IQ data from the Wechsler Intelligence Scale for Children and Wechsler Abbreviated Scale of Intelligence. The SSC cohort included 2,623 individuals with ASD from 2,623 families. The ASD simplex family participants in the Simons Foundation Autism Research Initiative (Fischbach, & Lord, 2010) had fewer defined medical conditions than the AGRE cohorts. Thus, for SSC, we selected the medical conditions defined in AGRE for replication analyses (Table 1).

Variables

For ASD multiplex families (AGRE cohort), assessed demographic variables included child age, sex, race, and ethnicity (Table 1). Race and ethnicity was determined for all participants in the AGRE cohort (Fischbach, & Lord, 2010; Geschwind et al., 2001) by parent report. Fourteen categories of medical conditions selected from medical history data for 627 individuals

with ASD are presented in Table 1. Medical phenotype variables are dichotomous. Specific details of GID in the AGRE sample were reported previously (Campbell et al., 2009). Frequencies of specific GID and sleep problems reported in AGRE and SSC are provided in Supporting Information Figure S1. For seizures, limited or no details (e.g. seizure frequency or seizure type) were available for either cohort. Seizure presence was determined from the medical history file and item 65 of the ADI-R, which rates seizures. For the latter, a score of 0 (no attacks) was considered seizure negative, 1 (history of attacks that might be epileptic, but diagnosis not established) or 2 (definite diagnosis of epilepsy) was considered seizure positive. Individuals that reported a febrile seizure only, with no continuing daily medication outside of the period of fever (7 on the ADI-R) were excluded from analysis ($n = 15$).

The demographic variables for ASD simplex families (SSC cohort) included child age, sex, race, and ethnicity. Race and ethnicity was determined for all participants in the SSC cohort (Fischbach, & Lord, 2010). As noted above, variables that comprised the medical history data differed from those collected in the AGRE sample, with information focusing on conditions that can be both debilitating and common in children with ASD. Five

medical conditions, also present in the AGRE sample, were reported for a large number of the SSC families (2,623). Medical phenotype variables are dichotomous (yes/no or true/false). Values listed as "suspected" or "not-sure" were considered unknown and excluded from analysis. Again, participants with a reported febrile seizure without any additional seizures and mothers with reported seizures during pregnancy were excluded. To be consistent with data included in AGRE, in which allergies are a single category without consistent further information, an "allergies" variable was created for SSC by combining the responses for environmental allergies and food allergies. Participants who responded "no" for both environmental allergies and food allergies were considered to have allergies absent. Participants who responded "true" for environmental allergies, but "false" for environmental allergies animal, were considered "true" for environmental allergies animal if environmental allergies other listed cat, dog, or feathers. Combining the responses for having any history of sleep problems created the "sleep" variable.

The prevalence of 30 medical phenotypes was determined in the ASD multiplex family cohort. Fourteen co-occurring medical conditions were present in more than 10% of individuals and selected for further analysis. DSM-V (APA, 2013) includes sensory abnormalities as a component of the core features of ASD. Thus, though listed in AGRE as distinct medical conditions, they were not included in the cluster analyses. Eleven additional phenotypes occurred in less than 10% of the ASD cohort. These additional conditions included cerebral (5.0%), bone (5.5%), joint (6.7%), structural eye (7.5%) or ear (8.9%) abnormalities, craniofacial (5.5%), heart (4.8%), kidney and urinary (3.6%), genital (7.8%), endocrine growth (2.0%), or hearing (6.3%) problems. These low occurrence conditions were not considered in additional analyses. No association between medication use and the presence of GID was found in our previous prospective study of children with ASD (Gorrindo et al., 2012). In addition, because a temporal link between medication use and medical symptoms cannot be made in a retrospective study, medication was not examined as a covariate in this study.

Parent-reported SRS T-scores for their children in the AGRE or SSC cohorts were used as a measure of social and communication symptom severity. Vineland Composite, Communication, Social and Daily Living Skills Standard scores were used for domain analysis of impairments in communication, sociability and adaptive behaviors.

Statistical Analysis

Prevalence of comorbidities. Prevalence of participant characteristics and medical conditions were computed for the AGRE and SSC cohorts. Within the AGRE sample, there was a subsample ($N = 627$) with complete comorbidity data used for cluster analyses. Participant characteristics and comorbidity prevalence were compared between those included and excluded using generalized estimating equations to account for familial relationships. This analysis was done in order to determine whether the complete data subsample was representative of the larger AGRE sample. Models were linear for continuous characteristics, and logistic for dichotomous characteristics.

Additionally, we computed prevalence of medical conditions in unaffected family members of the children with ASD in the cohorts, in order to examine whether comorbidities were likely to be ASD-related, rather than due to familial risk. The AGRE ASD multiplex family sample consisted of 1,612 individuals from the same 351 families as the ASD multiplex family cohort of 728 individuals with ASD. For the AGRE cohort, data were available for allergies, asthma, gastrointestinal disorders, and seizures for family members. For the SSC cohort, data on asthma and seizures was available for family members. Comparison of the prevalence of these comorbidities between family members with ASD and those without ASD was performed using generalized estimating equations, revealing repeated observations within families.

Cluster analysis. A cluster analysis was performed using AGRE participants with complete data for the 14 medical variables ($n = 627$). The analysis was done to determine whether there were subsets of participants with similar co-occurring medical conditions. Given the dichotomous nature of the data, which can prove problematic to the development of clusters, we used five random sorts to examine whether the same solution was achieved when the cases were sorted in different orders. Such sorts reduce the possibility of identification of cluster solutions by chance.

Secondary analysis of medical conditions. There was general inability to discriminate clusters using the analyses in the AGRE sample (see below). A secondary analysis was pursued to determine the prevalence of a subset of co-occurring medical conditions (GID, sleep problems and seizures) based on (1) being present in both AGRE and SSC; (2) GID and seizures being enriched in family members of the probands in AGRE; and (3) these conditions also having substantial medical impact and prevalence in children with ASD (Coury, Jones, Klatka, Winklosky, & Perrin, 2009). In order to determine whether there was increased risk for having different combinations of the three conditions, the prevalence of bivariate and trivariate combinations of the three disorders were computed. The

odds of having one of the disorders, given that a child was diagnosed with another, were assessed. Within AGRE, logistic generalized estimating equations were used. This method accounts for a lack of independence for observations within families. In SSC, in which only one child per family was diagnosed with ASD, logistic regression was used. Due to the possibility of interactions across the comorbidities, further analyses of the bivariate combinations were performed while stratifying the third variable. These models were adjusted for gender differences. Within AGRE, sleep data were missing for 77 participants, bringing the sample number for these analyses to 651 ASD-affected children with complete data for all three conditions. For the SSC replication sample, there were 2392 ASD-affected children with complete data for all three conditions.

Association of comorbid high-risk conditions with ASD symptom severity. Mixed effects models were created to test the impact of GID, sleep and seizure comorbidities on ASD symptom severity, represented by SRS T-score and Vineland Standard Scores in four domains. Two models were developed for SRS data: (1) the total number of comorbidities, ranging from 0–3, regardless of type, and (2) an additive model of the specific type of comorbidities. Models were adjusted for gender and AGRE cohort. In both AGRE and SSC, a very limited number of probands in our analysis of medical conditions had IQ data, precluding further analysis.

All statistical tests assumed a two-sided alternative hypothesis, a 0.05 significance level, and were performed using SPSS v.21 (IBM, Somers, NY).

Results
AGRE Cohort Characteristics

The ASD multiplex family cohort (Table 1) had a mean age of 9 ± 5 years, was predominantly male (78.3%) and of European (83%) and non-Hispanic (85%) racial and ethnic origins, respectively. There were 334 families represented by at least two ASD-affected children in the dataset. There were 627 ASD-affected children from 312 families with a complete medical history for the 14 most frequent co-occurring medical conditions identified in the ASD multiplex family cohort selected for use in the cluster analysis (Table 1). These children were similar, though not identical, to the larger ASD multiplex family cohort (see below). The prevalence of medical conditions ranged from 10.7% (abnormal growth patterns) to 77.4% (stereotypies). The prevalence of medical variables included in the final sample for cluster analysis was very similar to the entire AGRE sample, with small differences for a subset (Table 1).

Clustering of Co-Occurring Medical Conditions in Individuals with ASD from Multiplex Families

Cluster analysis of the 14 comorbidities showed unstable results that depended greatly on how the data were sorted. This indicates that while clusters could be identified, these were not sufficiently distinct to provide the specificity for subgrouping of children with ASD. The number of clusters across five random sorts was: 2, 4, 3, 2, 5, with all fits being poor. The analysis indicates that symptom co-occurrence analyzed across a large number of reported medical conditions in children with ASD does not provide sufficient specificity for determining unique components, a finding also reported in an analysis of hospital medical records (Doshi-Velez et al., 2014).

Co-occurring Medical Conditions in ASD Multiplex and Simplex Family Members

Medical conditions of ASD-unaffected family members typically are not reported in the analysis of comorbidities in probands diagnosed with ASD (Campbell et al., 2009; McElhanon et al., 2014). The large AGRE and SSC family cohorts provided an opportunity to evaluate this, given the parent-reported medical history data. To determine whether there was enrichment in family members of specific co-occurring medical conditions that are present in the probands, four medical conditions (allergies, asthma, GID, seizures) with data available for both ASD children and unaffected family members in the AGRE cohort were analyzed. Information on sleep problems was not available for unaffected family members, but was prevalent in the ASD probands and affected siblings. Table 2 reports the prevalence of co-occurring medical conditions in the multiplex and simplex ASD children, parents, and unaffected, non-ASD siblings. Within the AGRE sample, prevalence of allergies and asthma were similar between children with ASD and their parents and siblings, while GID and seizures were much more prevalent in children with ASD than in their unaffected family members. Information on allergies and GID in family members was not available for unaffected family members in SSC. The prevalence of asthma was similar across SSC family members, whereas the prevalence of seizures was higher in children with ASD.

For probands who report allergies, asthma, GID or seizures, we calculated the proportion of their family members with the same medical conditions (Table 2). Calculated odds ratios were adjusted for repeated measure within families. Having a family member reporting allergies increased the chance that probands had allergies, with the risk being highest if ASD-unaffected siblings had allergies (OR = 2.6; for all measures, p< .01). Odds ratios were much greater for asthma, ranging

Table 2. Prevalence of Medical Conditions in Families

	AGRE cohort						SSC cohort					
	Children with ASD		Parents		Non-ASD Siblings		Children with ASD		Parents		Non-ASD Siblings	
N	728		694		184		2623		5246		2623	
Allergies	291	40.0%	270	39.9%	56	30.4%	918	37.7%	—	—	—	—
Asthma	113	16.3%	104	15.2%	33	18.2%	259	10.6%	426	8.7%	267	11.3%
GID	309	42.4%	184	26.5%	23	12.2%	1121	43.1%	—	—	—	—
Seizures	89	12.2%	28	4.1%	5	2.7%	129	5.30%	45	0.9%	24	1.1%

Note. Percentages reported exclude unknown diagnoses.

from 3.1 to 3.9 if the condition was also reported in parents. Asthma was especially prevalent within ASD-unaffected siblings. Only 11% of probands with ASD had asthma, but of those children, a striking 73% had an unaffected sibling with asthma (OR = 22.1). There

Figure 1. Odds ratios of co-occurring medical symptoms. Graph depicts, for the AGRE and SCC cohorts, the odds ratio (adjusted for proband sex) of bivariate and trivariate combinations of a child having co-occurring gastrointestinal disturbance (GID), sleep disturbance, and seizures. Table 3 reports the specific ranges of calculated ORs and results of significance tests. Note that for some combinations in the AGRE sample, such as those including seizures in the GID x sleep multivariate analysis, the confidence intervals are large, reflecting small sample numbers in the seizure category.

was a statistically significant association of GID between probands and mothers (OR = 1.6, $P < 0.01$) and probands and unaffected siblings (OR = 2.2, $P < 0.006$), but not with probands and fathers. Seizure data in the two samples were sparse, with very few family members reporting seizures. There was a positive association, but this is interpreted with caution because of the limitations imposed by small sample sizes.

Co-Occurring GID, Seizure, and Sleep Problems in ASD

We performed an analysis of the occurrence of ASD with comorbid medical conditions reported in children in both the AGRE and SSC cohorts (Table 2). Within AGRE, 25% of children with ASD had none of the conditions, while 42% had one condition, 28% had two conditions, and 4.5% had all three conditions. The most common overlapping conditions with ASD were GID and sleep problems, which were found together in 23% of the sample. The SSC sample followed similar patterns of prevalence, though given the lower prevalence of seizures in this sample, there were fewer children with all three conditions (2.6%).

In both AGRE and SSC, there was a statistically significant association of GID with sleep or seizures, and no statistically significant association between sleep and seizures in the logistic models containing all three outcomes. Figure 1 shows the Odds Ratios (OR) stratified by various variables to depict the bivariate and trivariate effects (values and statistics reported in Table 3). For AGRE probands, a child with ASD and GID had an approximately two-fold increased risk of reporting sleep disturbance, seizures, or both. Within SSC, the OR of having the third comorbidity in the context of the presence of the other two was 1.7 for seizures, 2.0 for sleep disorders, and 1.9 for GID.

Relationship Between Co-Occurring Medical Conditions and Behavioral Symptom Severity

Models predicting social, communication, and adaptive symptom severity examined whether (a) the total number or (b) the type of medical condition the child
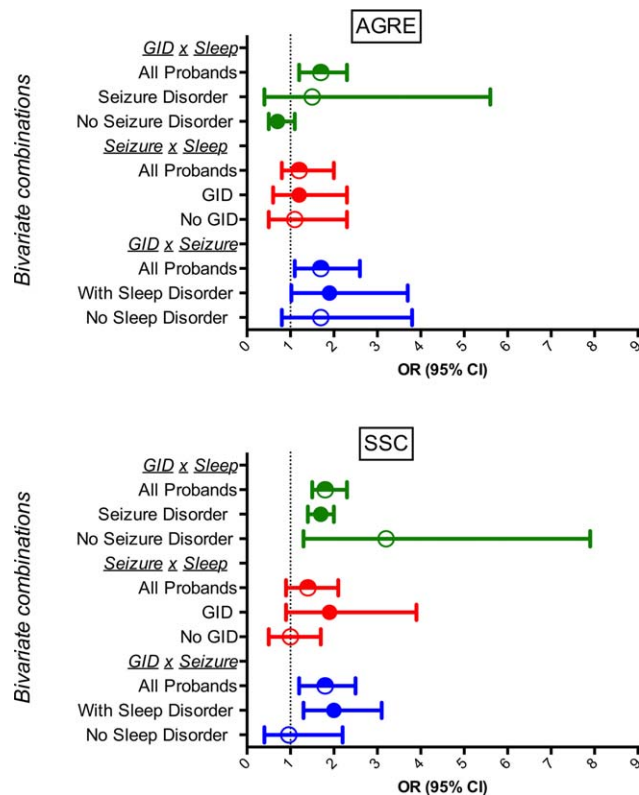
Table 3. Odds Ratios of Co-Occurring Medical Symptoms

	AGRE[a]				SSC[a]			
		95% CI				95% CI		
	OR	Lower	Upper	P	OR	Lower	Upper	P
GID & Sleep								
All probands (regardless of seizure status)	1.7	1.2	2.3	0.003	1.8	1.5	2.3	0.003
With seizure disorder	1.5	0.4	5.6	0.53	1.7	1.4	2.0	0.01
W/O seizure disorder	0.7	0.5	1.1	0.08	3.2	1.3	7.9	<0.001
Seizure × Sleep								
All probands (regardless of GID status)	1.2	0.8	2.0	0.41	1.4	0.9	2.1	0.13
GID	1.2	0.6	2.3	0.58	1.9	0.9	3.9	0.08
No GID	1.1	0.5	2.3	0.87	1.0	0.5	1.7	0.86
GID × Seizure								
All probands (regardless of sleep status)	1.7	1.1	2.6	0.02	1.8	1.2	2.5	0.002
W/ Sleep disorder	1.9	1.02	3.7	0.04	2.0	1.3	3.1	0.001
W/o sleep disorder	1.7	0.8	3.8	0.16	1.0	0.4	2.2	0.93

[a]Adjusted for sex & limited to participants with all three measured; AGRE adjusted for multiple probands per family.

exhibited impacted functional outcomes, adjusting for the child's sex. These competing models accounted for equal amounts of variability in the outcomes (6.7% versus 6.8% of the Vineland and 7.1% versus 7.2% of the SRS in the AGRE sample). Given these results, the more parsimonious model of total number of comorbidities was used. For the VABS, the presence of each additional medical condition lowered the composite score by a mean of 6.3 points ($P < 0.001$). Thus, there was a striking difference of approximately two standard deviations between children with ASD and no co-occurring medical conditions versus those with all three conditions. For the SRS, each additional condition increased the SRS by 3.6 points ($P < 0.001$). For the SSC cohort, the Vineland Composite score decreased by a mean of 2.7 points for each additional comorbidity ($P < 0.001$). The SRS exhibited a marginal decrease of 1.9 points for each additional comorbidity ($P < 0.04$).

DISCUSSION

This study had two major goals: (1) to determine whether there are relations among subsets of co-occurring medical conditions that may suggest a shared biologically-based risk in subsets of children with ASD; and (2) to determine whether individual medical conditions predict risk for the presence or development of other medical symptoms. We found that the prevalence of certain medical conditions was consistent in two genetically distinct ASD family cohorts (multiplex and simplex), specifically allergies, GID and seizures. Most important, the proband data in both AGRE and SSC cohorts revealed predictive relations among GID, seizures and sleep problems, three commonly reported and clinically challenging medical conditions for children

with ASD. The data from the original and replication cohorts show that clinical diagnosis of GID or sleep disturbances increases the risk for the presence of the other condition, as well as for seizures.

In our analysis of the prevalence of medical conditions in family members of the ASD-affected child, some conditions in ASD-unaffected siblings and parents may contribute modestly to predicting certain proband medical conditions. For example, the odds ratio for the presence of asthma and allergies in siblings was dramatically increased. The prevalence of both asthma and allergies was also similar in the ASD children and their family members, suggesting that these conditions are following their typical familial pattern and are less specifically related to ASD diagnosis. In contrast, GID was also overrepresented in unaffected siblings and mothers, though with a much smaller effect size for utility of predicting status of the proband and with a much higher prevalence in children with ASD than in their relatives. We interpret medical conditions in family members with caution because they were reported in relatively small numbers of ASD-unaffected relatives compared with the probands. Nonetheless, the analysis suggests that genetic and environmental factors shared among family members, distinct from the factors responsible for an ASD diagnosis, influence the expression of these medical conditions in specific children. Future prospective studies are needed to address this question more directly.

This study used several multivariate statistical approaches to address the primary goal of defining medical subgroups of co-occurring conditions that had both the sensitivity and specificity to cluster children with ASD. Beginning with the 14 most prevalent conditions reported in the AGRE sample, and reducing complexity by combining or eliminating very common

symptoms, none of the analyses revealed clinically useful clusters. A recent study (Doshi-Velez et al., 2014), using ICD-9 codes in medical records, reported a similar outcome, with a failure to cluster meaningful subsets of medical symptoms in 91.6% of their sample, with seizures, multisystem and psychiatric diagnoses overrepresented in 2.5%, 4%, and 4.1% of the children in the medical records with an ASD diagnosis, respectively. Thus, from the perspective of clinical utility, the data from the current and previous studies suggest that specific co-occurring symptoms do increase the likelihood of additional comorbidities, but that these do not form discreet clusters that typically group together within a substantial subset of the ASD population.

Secondary analysis of three of the medical symptoms reported in both AGRE and SSC probands revealed predictive co-occurring patterns. GID, sleep problems, and seizures are a challenge to manage clinically and can have a dramatic impact on quality of child and family life (Kohane et al., 2012; Lajonchere and Consortium, 2010). Moreover, for each of the three, correlations with ASD symptom severity have been reported (Gorrindo et al., 2012; Malow et al., 2012; Viscidi et al., 2013). Moreover, GID and seizures are enriched in family members in the AGRE sample. The analyses of the AGRE and larger SSC samples revealed a statistically significant relation between the co-occurrence of GID with either sleep problems, seizures, or both. Neither sleep problems nor seizures predicted the presence of another comorbidity in AGRE. However, in the larger SSC collection, GID or sleep problems individually increased the odds of the presence of the other or of all three medical conditions in the same child. Seizure presence was not predictive of the other symptoms in either cohort, but this may be due to the limited numbers of seizures reported in probands in AGRE and SSC. The increased power afforded by the larger SSC sample may be a factor in the emergence of both GID and sleep disturbances as predictors of comorbidities. We cannot rule out, however, that there are possible biological differences between children with ASD in single incident (SSC) and multi-incident (AGRE) families. Using data from both cohorts, the increased odds ratios for co-occurrence of specific symptoms indicate that clinicians should carefully assess children with one identified comorbid medical condition for the presence of others. Most specifically, children with identified GID or sleep disturbance (or both) should be screened for possible seizures.

The presence of GID, seizures and sleep problems together predicted more severe adaptive behavioral symptoms relevant to their ASD diagnosis, consistent with studies of children in which a functional GID was the medical focus (Gorrindo et al., 2012; Peters et al., 2013). In general, we found that medical comorbidities

had a significant impact on SRS and VABS-measured social, communication and adaptive behavioral skills in the multiplex AGRE cohort. In SSC families, the VABS-measured symptoms also worsened with increasing number of medical conditions, though the impact was more modest in this dataset. Surprisingly, in the SSC sample, the SRS score was lower for children with increasing number of medical conditions; although the difference reflects a small effect. It is not immediately obvious why this would be the case, but it could potentially indicate that medical conditions have a different relationship with primary ASD symptoms in children with less familial loading for ASD. Further work in cohorts containing both multiple and simplex families could clarify this finding. We also note that proband IQ may modify medical comorbidity risk. However, we were unable to analyze IQ as a covariate, because of the 728 probands in the AGRE sample with medical history data, only 18 of the 728 probands in the AGRE sample with medical history data also have composite IQ scores (SBI). Of the 2,623 probands in the SSC sample, 112 have composite IQ scores (WISC-IV or WASI) in the current analyses. However, our analysis of the relation between the VABS scores for adaptive behavior and the number of medical symptoms indicates that IQ is likely to be a relevant factor.

The results of the multivariate analyses were consistent between AGRE and SSC. These similarities contrast the genetic architectures present in the family-based and sporadic ASD samples examined in this study. These cohorts may have distinct genetic contributions that influence medical condition coexpression. Alternatively, medical condition coexpression may be modified by common, heritable variation, rather than single gene mutations in ASD more generally. Common variation was recently shown to underlie more than 50% of ASD diagnoses (Gaugler et al., 2014) and a common promoter variant that regulates transcription of the c-MET receptor tyrosine kinase is enriched in children with ASD and GID in AGRE multiplex families (Campbell et al., 2009). The effect of common variation on medical comorbidity expression will need to be examined directly in a much larger sample.

There are several important limitations to the datasets used for the present analyses. First, the AGRE dataset has broad representation of medical symptoms, which provided the opportunity for cluster analyses. The stratification resulted in some categories having limited sample sizes (even with eliminating those with prevalence less than 10%). Though the AGRE sample size was small, we note that a much larger study (Doshi-Velez et al., 2014)similarly failed to identify medical clusters for >90% of the sample (Doshi-Velez et al., 2014) . The SSC dataset included a limited number of medical symptoms. These were included in the

secondary analyses. Second, we note that while the multivariate analysis resolves potentially important relations, the sample size for seizures in particular is very limited. Thus, in some instances, the confidence intervals are quite large (e.g. GID x Sleep w/seizures), and speak to the need for studies in the future focusing on cohorts with greater representation of specific symptoms categories. Third, there is limited data on mental health disorders in family members in both AGRE and SSC, which precluded what would be an interesting analysis of possible enrichment in multiplex and/or simplex families. Fourth, clinical conditions for both the AGRE and SSC collections were obtained from medical history data derived from parent questionnaires, which may introduce a bias distinct from estimates of comorbidities obtained from hospital records. We note, however, that at least for GID, in our prospective study of children with a functional GID, there was greater than 90% agreement between a research validated parent questionnaire and diagnosis by a pediatric GI specialist for the presence of a gut disturbance [mostly constipation/diarrhea; (Gorrindo et al., 2012)]. Fifth, for sleep disturbances in children with ASD, it is believed that there is both underreporting of problems by parents and limited screening by clinicians (Malow et al., 2012). Thus, while there have been calls for broad-based screens for sleep and GID in ASD (Buie et al., 2010; Malow et al., 2012), the data here suggest that there is sufficiently increased risk for co-occurrence of the two such that the appearance of one should trigger assessment for the other. Sixth, seizure disorders appear more commonly in older children in ASD (Viscidi et al., 2013), which may occur later than symptom appearance of GID or sleep problems. Thus, together with identification of children with genetic mutations that have high risk for epilepsy (Jeste, & Geschwind, 2014), co-occurrence of both GID and sleep problems in the same child may justify more detailed neurological assessments (Jeste, & Geschwind, 2014). Additionally, our findings corroborate other reports of increased symptom severity in children with GID or seizures (Gorrindo et al., 2012; Jeste, & Geschwind, 2014). Together with an increased risk for sleep disturbances, these co-occurring medical conditions warrant intervention strategies that routinely use interdisciplinary approaches in which social communication and behavioral and medical symptoms are treated and evaluated for improvements in an integrated fashion.

A shared biology between a neurodevelopmental disorder and additional medical conditions is not unique to ASD. The recognition of specific patterns of neurodevelopmental dysfunction together with additional medical symptoms has enumerated many syndromes, which has facilitated gene identification and provided biological insight in many cases. However, the specific mechanism that leads to GID, sleep and seizures occurring within the same child also diagnosed with ASD is not yet known. Others and we have speculated that a developmental trajectory of shared biological risk of brain and relevant peripheral organ development may underlie comorbidity (Gorrindo et al., 2012; Isaksen et al., 2012; Laurier et al., 2015; Moeschler, & Shevell, 2014). This may reflect the pleiotropic nature of the gene disruptions that can cause specific developmental perturbations outside of the brain. Comorbidity of GID, sleep and seizures in girls with Rett Syndrome or children with Fragile X syndrome is well-reported (Jeste, & Geschwind, 2014), and Porges and colleagues have articulated a model of autonomic dysregulation (Heilman, Harden, Zageris, Berry-Kravis, & Porges, 2011; Porges, & Furman, 2010). Brainstem/autonomic involvement (Gorrindo et al., 2012) would be consistent with the enrichment of co-occurrence of GID, sleep problems and seizure disorders in ASD. Measuring central autonomic and integrative arousal systems in children with ASD may be valuable in addressing an autonomic/brainstem hypothesis of ASD and specific medical comorbidities, but to date there has been little work in this area. Prospective measures of central and peripheral autonomic function, done in the context of medical symptom assessments, could be designed for specific hypothesis testing. In relation to the developing brain, there are comorbidities that indicate broader disturbances in the development and functioning of circuits that are not specific to an ASD diagnosis (Geschwind, 2009; Kohane et al., 2012; Sacco et al., 2010). Perhaps underappreciated until recently, neurological disturbances, such as sleep problems (Marcano-Reik, Prasad, Weiner, & Blumberg, 2010) or seizures (Swann, 2004) can have profound effects on the developing brain. Moreover, there is recent recognition that the gut microbiome, in addition to impacting gut physiology by immune or autonomic dysfunction, can affect both brain and behavior (Collins, Surette, & Bercik, 2012; Cryan and Dinan, 2012). Together with mounting evidence from previous studies, the analysis here of two independent cohorts indicates that co-occurring medical conditions are individually common, frequently occur together, and are associated with greater impairment. Carefully screening for and targeting these medical conditions using a multidisciplinary treatment approach may improve not only the medical conditions themselves but also may improve the trajectory of ASD and cognitive symptoms by removing symptoms that interfere with response to behavioral and other interventions.

Acknowledgements

References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders, 5th edition: DSM-5. Arlington, VA: American Psychiatric Publishing.

Berg, J.M., & Geschwind, D.H. (2012). Autism genetics: searching for specificity and convergence. Genome Biology 13: 247.

Boomsma, A., Van Lang, N.D., De Jonge, M.V., De Bildt, A.A., Van Engeland, H., & Minderaa, R.B. (2008). A new symptom model for autism cross-validated in an independent sample. Journal of Child Psychology and Psychiatry 49:809-816.

Bradford, Y., Haines, J., Hutcheson, H., Gardiner, M., Braun, T., Sheffield, V., et al (2001) Incorporating language phenotypes strengthens evidence of linkage to autism. American Journal of Medical Genetics 105:539-547.

Buie, T., Campbell, D.B., Fuchs, G.J., 3rd, Furuta, G.T., Levy, J., & Vandewater, J. (2010). Evaluation, diagnosis, and treatment of gastrointestinal disorders in individuals with ASDs: a consensus report. Pediatrics 125 Suppl 1:S1-18.

Campbell, D.B., Buie, T.M., Winter, H., Bauman, M., Sutcliffe, J.S., & Perrin, J.M. (2009). Distinct genetic risk based on association of MET in families with co-occurring autism and gastrointestinal conditions. Pediatrics 123:1018-1024.

Collins, S.M., Surette, M., & Bercik, P. (2012). The interplay between the intestinal microbiota and the brain. Nature Reviews Microbiology 10:735-742.

Constantino, J.N. (2002). The social responsiveness scale. Los Angeles, CA: Western Psychological Services.

Constantino, J.N., Gruber, C.P., Davis, S., Hayes, S., Passanante, N., & Przybeck, T. (2004). The factor structure of autistic traits. Journal of Child Psychology and Psychiatry 45:719-726.

Coury, D., Jones, N.E., Klatka, K., Winklosky, B., & Perrin, J.M. (2009). Healthcare for children with autism: the Autism Treatment Network. Current Opinion in Pediatrics 21:828-832.

Cryan, J.F., & Dinan, T.G. (2012). Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. Nature Reviews Neuroscience 13:701-712.

Doshi-Velez, F., Ge, Y., & Kohane, I. (2014). Comorbidity clusters in autism spectrum disorders: an electronic health record time-series analysis. Pediatrics 133:e54-63.

Fischbach, G.D., & Lord, C. (2010). The Simons Simplex Collection: a resource for identification of autism genetic risk factors. Neuron 68:192-195.

Frazier, T.W., Youngstrom, E.A., Kubu, C.S., Sinclair, L., & Rezai, A. (2008). Exploratory and confirmatory factor analysis of the autism diagnostic interview-revised. Journal of Autism and Developmental Disorders 38:474-480.

Gaugler, T., Klei, L., Sanders, S.J., Bodea, C.A., Goldberg, A.P., Lee, A.B., et al. (2014). Most genetic risk for autism resides with common variation. Nature Genetics 46:881-885.

Geschwind, D.H. (2009). Advances in autism. Annual Review of Medicine 60:367-380.

Geschwind, D.H., Sowinski, J., Lord, C., Iversen, P., Shestack, J., Jones, P., et al. (2001). The autism genetic resource exchange: a resource for the study of autism and related neuropsychiatric conditions. The American Journal of Human Genetics 69:463-466.

Gorrindo, P., Williams, K.C., Lee, E.B., Walker, L.S., McGrew, S.G., Levitt, P. (2012). Gastrointestinal dysfunction in autism: parental report, clinical evaluation, and associated factors. Autism Resarch 5:101-108.

Heilman, K.J., Harden, E.R., Zageris, D.M., Berry-Kravis, E., & Porges, S.W. (2011). Autonomic regulation in fragile X syndrome. Developmental Psychobiology 53:785-795.

Hilton, C.L., Zhang, Y., Whilte, M.R., Klohr, C.L., & Constantino, J. (2012). Motor impairment in sibling pairs concordant and discordant for autism spectrum disorders. Autism 16:430-441.

Isaksen, J., Bryn, V., Diseth, T.H., Heiberg, A., Schjoberg, S., & Skjeldal, O.H. (2012). Children with autism spectrum disorders–the importance of medical investigations. European Journal of Paediatric Neurology 17:68-76.

Jeste, S.S., & Geschwind, D.H. (2014). Disentangling the heterogeneity of autism spectrum disorder through genetic findings. Nature Reviews Neurology 10:74-81.

Kohane, I.S., McMurry, A., Weber, G., MacFadden, D., Rappaport, L., Kunkel, L., et al. (2012). The co-morbidity burden of children and young adults with autism spectrum disorders. PLoS ONE 7:e33224.

Lajonchere, C.M., AGRE Consortium. (2010). Changing the landscape of autism research: the autism genetic resource exchange. Neuron 68:187-191.

Laurier, V., Lapeyrade, A., Copet, P., Demeer, G., Silvie, M., Bieth, E., et al. (2015). Medical, psychological and social features in a large cohort of adults with Prader-Willi syndrome: experience from a dedicated centre in France. Journal of Intellectual Disability Resarch. 59:411-421.

Lecavalier, L. (2006). Behavioral and emotional problems in young people with pervasive developmental disorders: relative prevalence, effects of subject characteristics, and empirical classification. Journal of Autism and Developmental Disorders 36:1101-1114.

Lord, C., Rutter, M., DiLavore, P., & Risi, S. (1999). Autism diagnostic observation schedule manual. Los Angeles, CA: Western Psychological Services.

Malow, B.A., Byars, K., Johnson, K., Weiss, S., Bernal, P., Goldman, S.E., et al. (2012). A practice pathway for the identification, evaluation, and management of insomnia in children and adolescents with autism spectrum disorders. Pediatrics 130 Suppl 2:S106-124.

Marcano-Reik, A.J., Prasad, T., Weiner, J.A., & Blumberg, M.S. (2010). An abrupt developmental shift in callosal modulation of sleep-related spindle bursts coincides with the emergence of excitatory-inhibitory balance and a reduction of somatosensory cortical plasticity. Behavioral Neuroscience 124:600-611.

McElhanon, B.O., McCracken, C., Karpen, S., & Sharp, W.G. (2014). Gastrointestinal symptoms in autism spectrum disorder: a meta-analysis. Pediatrics 133:872-883.

Memari, A., Ziaee, V., Mirfazeli, F., & Kordi, R. (2013). Investigation of autism comorbidities and associations in a school-based community sample. Journal of Child Adolescent Psychiatric Nursing 25:84-90.

Moeschler, J.B., & Shevell, M. (2014). Comprehensive evaluation of the child with intellectual disability or global developmental delays. Pediatrics 134:e903-e918.

Murdoch, J.D., & State, M.W. (2013). Recent developments in the genetics of autism spectrum disorders. Current Opinion in Genetics & Development 23:310-315.

Peters, J.M., Taquet, M., Vega, C., Jeste, S.S., Fernandez, I.S., Tan, J. et al. (2013). Brain functional networks in syndromic and non-syndromic autism: a graph theoretical study of EEG connectivity. BMC Medicine 11:54.

Porges, S.W., & Furman, S.A. (2010). The early development of the autonomic nervous system provides a neural platform for social behaviour: a polyvagal perspective. Infant and Child Development 20:106-118.

Rutter, M., Le Couteur, A., & Lord, C. (2003). Autism diagnostic inventory–Revised. Los Angeles, CA: Western Psychological Services.

Sacco, R., Curatolo, P., Manzi, B., Militerni, R., Bravaccio, C., Frolli, A., et al. (2010). Principal pathogenetic components and biological endophenotypes in autism spectrum disorders. Autism Research 3:237-252.

Scherer, S.W., & Dawson, G. (2011). Risk factors for autism: translating genomic discoveries into diagnostics. Human Genetics 130:123-148.

Sikora, D.M., Johnson, K., Clemons, T., & Katz, T. (2012). The relationship between sleep problems and daytime behavior in children of different ages with autism spectrum disorders. Pediatrics 130 Suppl 2:S83-90.

Silverman, J.M., Smith, C.J., Schmeidler, J., Hollander, E., Lawlor, B.A., Fitzgerald, M., et al. (2002). Symptom domains in autism and related conditions: evidence for familiality. American Journal of Medical Genetics 114:64-73.

Sparrow, S., Cicchetti, D., & Balla, D. (2005). Vineland adaptive behavioral scales, 2nd edition. Minneapolis, MN: Pearson Assessment.

Swann, J.W. (2004). The effects of seizures on the connectivity and circuitry of the developing brain. Mental Retardation and Developmental Disabilities Research Reviews 10:96-100.

Tadevosyan-Leyfer, O., Dowd, M., Mankoski, R., Winklosky, B., Putnam, S., McGrath, L., et al. (2003). A principal components analysis of the Autism Diagnostic Interview-Revised. Journal of the American Academy of Child and Adolescent Psychiatry 42:864-872.

Veatch, O.J., Veenstra-Vanderweele, J., Potter, M., Pericak-Vance, M.A., & Haines, J.L. (2014). Genetically meaningful phenotypic subgroups in autism spectrum disorders. Genes Brain and Behavior 13:276-285.

Viscidi, E.W., Triche, E.W., Pescosolido, M.F., McLean, R.L., Joseph, R.M., Spence, S.J., et al. (2013). Clinical characteristics of children with autism spectrum disorder and co-occurring epilepsy. PLoS One 8:e67797.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Frequencies of specific GID and sleep problems reported in children with ASD in the AGRE and SSC cohorts. Specific gastrointestinal dysfunctions (GID) were reported for 309 children with ASD in the AGRE cohort and 2,622 children with ASD in the SSC cohort. Constipation is the most frequent GID reported in both cohorts. Specific sleep problems were reported for 355 children with ASD in the AGRE cohort and 2,622 children with ASD in the SSC cohort. All variables were reported as present (black) or absent (grey), except for Sleep Disorder Course (*) in the AGRE cohort, which was reported as continuous (black) or intermittent (grey).