Structural magnetic resonance imaging data do not help support DSM-5 Autism Spectrum Disorder category

Laura Pina-Camacho, Sonia Villero, Leticia Boada, David Fraguas, Joost Janssen, Maria Mayoral, Cloe Llorente, Celso Arango, Mara Parellada

Child and Adolescent Psychiatry Department, CIBERSAM, Instituto de Investigación Sanitaria Gregorio Marañón, IISGM. Hospital General Universitario Gregorio Marañón. Ibiza 43, 28009 Madrid, Spain.

Unit of Child and Adolescent Mental Health, Department of Psychiatry, Complejo Hospitalario Mancha Centro. Avenida Constitucion, 3. 13600 Alcázar de San Juan. Ciudad Real, Spain.

Mental Health Department, CIBERSAM, Complejo Hospitalario Universitario de Albacete. Hermanos Falcó, 37. 02006 Albacete, Spain.

Corresponding author’s contact information: Laura Pina-Camacho, Child and Adolescent Psychiatry Department. CIBERSAM, Instituto de Investigación Sanitaria Gregorio Marañón, IISGM. Hospital General Universitario Gregorio Marañón. Ibiza 43, 28009 Madrid, Spain. Tel.: +34 91 426 50 05; Fax: +34 91 426 50 04; e-mail: lpina@iisgm.com

Abstract:
This systematic review aims to determine whether or not structural magnetic resonance imaging (sMRI) data support the DSM-5 proposal of an Autism Spectrum Disorder (ASD) diagnostic category, and whether or not classical DSM-IV Autistic Disorder (AD) and Asperger syndrome (AS) categories should be subsumed into it. The most replicated sMRI findings in patients with ASD compared with healthy controls are increased total brain volume in early childhood and decreased corpus callosum volume. Regarding the notion of a spectrum, some studies support that AS and AD are similar but “quantitatively different” diagnostic categories, whereas others support that they are “qualitatively different” entities with specific brain structural abnormalities. It seems that there are still not enough arguments from sMRI data for or against subsuming DSM-IV categories under a single ASD category.

Keywords: Autism spectrum disorder; autistic disorder; Asperger syndrome; structural magnetic resonance imaging; DSM-5
1 Introduction

In 1979, Wing launched the concept of the *autistic continuum or spectrum* (Wing & Gould, 1979) and, nine years later, Allen coined the term *autism spectrum disorders* (Allen, 1988). However, controversy still surrounds Wing’s original concept of a broad autistic phenotype. In fact, clinicians and researchers have used the term *Autism Spectrum Disorders* to include Autistic Disorder (AD) (including high functioning autism –HFA– and low functioning autism –LFA), Asperger Syndrome (AS), and Pervasive Developmental Disorder Not Otherwise Specified (PDD NOS) (Levy, Mandell, & Schultz, 2009), where the term “spectrum” reflects the variability in symptom severity among patients. In this context, there is a proposal for the forthcoming DSM-5 diagnostic classification to create the broad diagnostic category of Autism Spectrum Disorder (ASD) (see www.dsm5.org). Nevertheless, there are still unanswered questions about the ASD construct, which have led to the current debate on how ASD should best be conceptualized in DSM-5 (Frazier et al., 2012; Happe, 2011; Mandy, Charman, & Skuse, 2012; Mattila et al., 2011; Pina-Camacho et al., 2012; Tanguay, 2011). One of these questions is whether ASD constitutes a well defined biological entity compared with the earlier Pervasive Developmental Disorders, and whether classic DSM-IV categories – and especially AS – should be subsumed into this broader category.

Over the past few decades, magnetic resonance imaging (MRI), a non-invasive in vivo technique, has allowed access to the anatomy and physiology of the developing brain and has contributed to our understanding of neurodevelopment in health and illness (Giedd & Rapoport, 2010). In the late eighties, researchers started using structural MRI (sMRI) to examine pathological changes in the brain structure of pediatric and psychiatric patients (Mana, Paillere Martinot, & Martinot, 2010; Potts, Davidson, & Krishnan, 1993), including those with autism. Widely used sMRI techniques are summarized in Table 1. Initially, studies measured volume by totaling the amount of voxels in manually predefined regions of interest (ROIs). These methods were followed up by voxel-based approaches such as voxel-based morphometry (VBM) (Whitwell, 2009), which allows whole brain exploration of structural differences and thus does not depend on manually predefined regions. More recently, with the advent of improved image acquisition (e.g., higher field strength, higher isotropic voxel resolution, and improved gray-white matter contrast), additional morphometric measures that focus on the thickness, surface area, and curvature of the cortex have emerged. Furthermore, multivariate statistical analysis frameworks have been developed that aim to classify subjects as patients or controls based on large morphometric datasets (Chung, Bubenik, & Kim, 2009; Chung, Dalton, Shen, Evans, & Davidson, 2007; Chung, Robbins, &
There has been a steady rise in the number of sMRI publications both in the forthcoming ASD category and in the classic DSM-IV categories, mainly in AD and AS (Mana et al., 2010). In these studies, the authors have tried to define the neurological underpinnings of these diagnostic entities and to relate brain structural abnormalities with associated behavioral and clinical features. However, few studies have tried to summarize the findings of previous sMRI reports in order to find neuroanatomic evidence for the new ASD diagnostic category (Chen, Jiao, & Herskovits, 2011; Stigler, McDonald, Anand, Saykin, & McDougle, 2011; Verhoeven, De Cock, Lagae, & Sunaert, 2010). The objective of this review is to assess whether or not reported sMRI findings support the proposal of subsuming DSM-IV categories under this new ASD category based on specific neuroanatomical substrates.

2 Methods

We conducted a systematic Pubmed search on structural MRI studies of ASD published in English between January 1990 and February 2012. The following database search strategy was used: ‘("Autism spectrum disorders"[All Fields] OR "Asperger syndrome"[All Fields] OR "Asperger's syndrome"[All Fields] OR "Autistic disorder"[All Fields]) AND ("Magnetic resonance imaging"[All Fields]) NOT pubstatus ahead of print’. After excluding in-press papers, as not all of them were available in full text, 663 records were identified and screened. Of these, 256 full-text articles were eligible, as they fulfilled all the following inclusion criteria: a) being an original article or a review; b) including patients with ASD, autistic disorder (HFA or LFA), AS, PDD NOS, or childhood disintegrative disorder (CDD); and c) providing structural MRI data. We also identified 29 relevant studies that were referenced in these 257 eligible studies but did not appear in the initial database search. Thus, a total of 285 studies were finally included in this review. A total of 407 full-text articles were excluded because a) they were not a review or an original article (n=44); b) they did not focus on patients with ASD, AS, or AD (n=80); c) they did not provide neuroimaging data (n=48); or d) they used neuroimaging techniques other than sMRI, i.e., functional MRI, diffusion tensor imaging, positron emission tomography, or single-photon emission computed tomography (n=235). We decided not to include studies using neuroimaging techniques other than sMRI as that was not the main objective of this review and they have been reviewed elsewhere (Pina-Camacho et al., 2012). We also decided not to include Rett syndrome in this review, as it has disappeared from the DSM-5 proposal.
3 Results

Most studies included in this review compare patients with ASD (without distinction between DSM-IV subcategories) or HFA with healthy controls (HC). We found eight studies comparing patients with HFA and AS (Haznedar et al., 2006; Jou, Minshew, Keshavan, & Hardan, 2010; Kwon, Ow, Pedatella, Lotspeich, & Reiss, 2004; McAlonan et al., 2009; McAlonan et al., 2008; Toal et al., 2010; Via, Radua, Cardoner, Happe, & Mataix-Cols, 2011; Yu, Cheung, Chua, & McAlonan, 2011), four studies comparing patients with LFA, HFA, AS, and HC (Lotspeich et al., 2004; Nordahl et al., 2007; Scott, Schumann, Goodlin-Jones, & Amaral, 2009; Schumann et al., 2004), one study comparing patients with HFA, LFA, and HC (Salmond, Vargha-Khadem, Gadian, de Haan, & Baldeweg, 2007), and one study comparing patients with PDD-NOS, HFA, and HC (Lahuis et al., 2008). We did not find any study providing sMRI data on childhood disintegrative disorder (CDD). Firstly, we will summarize the main findings on structural abnormalities in patients with ASD and, secondly we will present those studies that focus on neuroanatomic differences between classic DSM subcategories, mainly between HFA and AS.

3.1 Structural MRI studies on ASD

Although research has highlighted the role of several specific brain regions in ASD pathogenesis (Volkmar & Pauls, 2003), the available literature on specific structural brain abnormalities presents discrepant results and may have been limited by methodological issues (Brambilla et al., 2003; Eliez & Reiss, 2000; Lord, Cook, Leventhal, & Amaral, 2000). The most replicated sMRI differences between ASD and age-matched HC were increased total brain volume (TBV) in early childhood, and increased cerebellar volume and decreased corpus callosum (CC) volume in patients with ASD. Volumetric studies on subcortical structures, such as the amygdalohippocampal system or thalamus, have shown inconsistent results. Finally, negative results (that is, non-significant volumetric differences) have also been reported between patients with ASD and HC. The main structural findings in patients with ASD compared with HC, classified by significant structures or regions, together with the main neuropathological and clinical correlates, whenever such data are available, have been provided separately in Supplementary Tables S1 and S2.
3.2 Structural MRI studies comparing categories derived from DSM-IV-TR: AS, HFA, LFA, and PDD NOS

Although these studies provide heterogeneous data in terms of structural abnormality patterns among diagnostic categories, we have classified them into two categories. Firstly, those studies that support the contention that LFA, HFA, and AS are similar but “quantitatively” different diagnostic entities, as they report similar volumetric abnormalities among these disorders along a quantitative continuum of severity; and secondly, those that consider that they are different diagnostic categories with “qualitatively” different abnormality patterns among them (as they show, for instance, different regional distributions of volumetric abnormalities). In all these studies, the diagnostic criteria for distinguishing individuals with AD (HFA and/or LFA) from those with AS, and for assigning groups, were defined *a priori*.

In the first category, there is a recent meta-analysis where the authors searched published VBM studies measuring gray matter (GM) volume in patients with ASD and HC and describing the proportion of participants with AD and AS in their samples (Via et al., 2011). Ten studies were included in the meta-analysis, and GM volume differences between diagnostic subgroups were analyzed. No significant differences were found in whole-brain or regional GM volume between the patients with AD (N=211) and the age-, sex-, and IQ-matched AS group (N=67). The authors concluded that both disorders may have similar neural substrates. However, the results were limited by the small number of studies included and by the varied diagnostic tools and criteria these studies used to assign individuals to the AD or the AS group (Via et al., 2011).

Another study reported that the mean brain GM volume for the AS group was intermediate between the HFA and HC groups, concluding that this may indicate a quantitative “continuum” in which brain GM volume increases with the severity of the ASD condition (Lotspeich et al., 2004). However, when performing correlations of specific brain volumes with estimated intelligence quotient (IQ) scores, there was a negative correlation between brain GM volume and performance IQ (pIQ) within the HFA but not the AS group, and a positive correlation between brain white matter (WM) volume and pIQ within the AS but not the HFA group. Thus, the authors concluded that these findings suggested qualitative differences in terms of neurodevelopment between subjects with HFA and AS. Regarding the LFA group, there was an unusually large variance in total brain tissue, suggesting that, neuroanatomically, it may represent a more heterogeneous population than the HFA or AS categories (Lotspeich et al., 2004).
Other studies reported quantitative differences between ASD categories within specific brain regions. Larger caudate volumes were detected in a group with AD, and intermediate volumes in an AS group, compared with HC (Haznedar et al., 2006). The sample in this study was small (N=17) and heterogeneous (IQ range=55-125). Furthermore, patients and HC were not IQ-matched and two patients had a history of exposure to antipsychotics (Haznedar et al., 2006). Another study reported a 16% larger amygdala in young children with –both LFA and HFA regardless of IQ, whereas in those with AS it was 9% larger (Schumann et al., 2004). Additionally, compared with an HC group, the hippocampus was 10% larger in the HFA group and slightly, but not significantly, larger in the AS group (Schumann et al., 2004).

The second group of sMRI studies reported qualitative brain volume differences (e.g., in terms of regional distribution) between patients with AS and HFA, thus concluding that they are qualitatively different diagnostic categories with specific neuroanatomical substrates, and so should not be subsumed into a single ASD category. As an example, there is a meta-analysis using a methodology similar to the one used by Via et al (Via et al., 2011) with a total of 18 studies comparing either patients with AD versus HC or patients with AS versus HC, using VBM methods and representing GM differences in a stereotactic space (Yu et al., 2011). Greater GM excess was reported in the AD group than in the AS group compared with HC. However, differences in the regional distribution of this GM excess were also detected. Whereas the AS group showed greater GM volume mainly in the left hemisphere, the AD group had a greater bilateral excess. Although this result could be considered a quantitative difference between the two disorders (greater GM excess in the AD group, with more marked changes in the left hemisphere), the authors concluded that this should be considered a qualitative difference because of the different regional distribution of this GM excess. An important limitation of this meta-analysis is that it did not directly compare the AS and AD groups (Yu et al., 2011). In parallel with these “hemisphere-distributed” GM volume differences between the two disorders, another study reported that abnormalities of WM systems affected mainly the left hemisphere in patients with HFA and the right hemisphere in those with AS (McAlonan et al., 2009).

Another study in male children and adolescents with LFA, HFA, AS, and HC, which conducted morphometric and geometric MRI analyses using surface-based morphometry, identified qualitatively different patterns of cortical folding abnormalities in LFA, HFA, and AS, particularly in children. Whereas LFA and HFA individuals had similar shape abnormalities centered on the pars opercularis of
the left inferior frontal gyrus, which were smaller and more posterior in the HFA group, the AS group showed deeper intraparietal sulci bilaterally (Nordahl et al., 2007).

In this second group of studies, the hypothesis that people with AS would have significant qualitative differences from those with AD in the anatomy of classic language regions has also been tested. Increased GM volume in the superior temporal, inferior parietal, and supramarginal gyri was reported in patients with AD, but not in those with AS compared with age and sex-matched HC (Toal et al., 2010). Additionally, one study found a higher left inferior gyrification index in a group with HFA compared to a AS and a HC group, being this increase more prominent in the inferior frontal regions, including Broca’s area (Jou et al., 2010). Moreover, a negative correlation between GM volume in Broca’s area in early childhood and age of language acquisition was described only in patients with HFA but not with AS (McAlonan et al., 2008). Different qualitative patterns of structural abnormality between ASD categories have also been described in other brain regions. For instance, vermal volume was found to be decreased in a sample of male children and adolescents with HFA compared with HC, but not in those with AS or LFA (Scott et al., 2009). However, the decrease in overall vermal volume was also observed in the ASD group as a whole compared with HC, and sample size was small among the three LFA, HFA, and AS phenotypes (n=62) (Scott et al., 2009). Another study reported divergent patterns of GM volume abnormalities in a group of children with “low ASD” (defined as those patients with verbal IQ < 85) and with “high ASD” (with verbal IQ > 85, including those with a clinical diagnosis of HFA and AS) (Salmond et al., 2007). The low ASD group showed decreased and the high ASD group increased cerebellar GM volume compared with HC. The high ASD group also had no significant volume abnormality in the postcentral gyrus and the dorsolateral prefrontal cortex, whereas the low ASD group had significantly decreased GM volume in these regions compared with HC. This study had a small sample size (n=44), did not differentiate between children with AS and HFA, and the results may have been confounded by the individual’s IQ (Salmond et al., 2007). Regarding limbic and subcortical structures, smaller GM volumes in the caudate and thalamus (McAlonan et al., 2008) and in the body of the cingulate gyrus (Kwon et al., 2004) were reported in patients with AS, but not with HFA, compared with HC. Additionally, patients with HFA showed smaller GM volume in pallidal-frontal regions compared with HC, not present in individuals with AS (McAlonan et al., 2008). Concerning the PDD NOS category, there is only one sMRI study comparing subjects with multiple complex developmental disorder (MCDD), subjects with HFA, and HC where the authors considered MCDD a phenotypically
defined subtype of PDD NOS (Lahuis et al., 2008). Similar brain structural abnormalities in subjects with MCDD and AD were reported, but without enlargement of head size in the former group, suggesting different developmental trajectories within these two subtypes of ASD categories (Lahuis et al., 2008).

4 Discussion

Despite of the extensive literature on the use of sMRI in patients with ASD, there are few consistent structural findings (Chen et al., 2011). Moreover, these findings do not always correlate with clinical, neuropathological, or neuropsychological features of the disorder (Eliez & Reiss, 2000; Griebling et al., 2010; Hardan, Girgis, Adams, et al., 2006; Hardan, Girgis, Lacerda, et al., 2006; Lord et al., 2000). Thus, it seems that sMRI data currently do not help resolve the DSM-IV versus DSM-5 controversy. In other words, there are still not enough arguments for or against subsuming DSM-IV categories under a single ASD category.

4.1 Structural MRI studies on ASD

Although there are several replicated findings when comparing patients with ASD and healthy individuals, such as increased TBV and cerebellar volume and decreased CC volume in the former group, there are also contradictory or even negative findings in terms of brain structural differences between the two groups.

4.2 AD and AS: similar or dissimilar disorders?

Regarding the DSM-5 proposal for a “spectrum” and the debate on whether AS and AD are “more dissimilar than similar” (Ozonoff, Rogers, & Pennington, 1991; Rinehart, Bradshaw, Brereton, & Tonge, 2002; Volkmar, Klin, Schultz, Rubin, & Bronen, 2000), sMRI results are heterogeneous. The same brain structure has been described as larger, smaller, or not different among the ASD clinical subgroups. For example, ventral temporal lobe GM volume has been found to be similar (McAlonan et al., 2008), increased (Yu et al., 2011), or decreased (Kwon et al., 2004) in AS and HFA compared with HC. Nevertheless, we can classify sMRI studies in two main groups, depending on whether or not they support the idea of a “spectrum” under a single ASD category. Among the former, the fact that volume abnormalities are similar in LFA, HFA, and AS groups, but usually more severe in LFA than in HFA, and in HFA than in AS, leads to the prevailing view of AS and AD as “qualitatively similar but quantitatively
different” diagnostic entities (Hardan et al., 2008) within a spectrum or continuum of severity. These findings are congruent with data from clinical and behavioral studies that do not find relevant differences between patients with AS and HFA (Howlin, 2003; Kaland, Mortensen, & Smith, 2007; Kamp-Becker et al., 2010; Klin, 2000; Manjiviona & Prior, 1995; Miller & Ozonoff, 2000; Szatmari et al., 2000), and support considering severity as a clinical specifier with previous subcategories subsumed under one single concept. On the other hand, another group of sMRI studies supports the concept that AS and AD are different diagnostic entities, with distinct and specific patterns of structural brain abnormalities. This would also be congruent with data showing qualitative differences between the two disorders in social skills (Gepner & Mestre, 2002; Rinehart, Bradshaw, Brereton, & Tonge, 2001; Szatmari, Archer, Fisman, Streiner, & Wilson, 1995), cognitive and executive functioning (Ehlers et al., 1997; Rinehart, Bradshaw, Moss, Brereton, & Tonge, 2000, 2001), clinical prognosis (Cederlund, Hagberg, Billstedt, Gillberg, & Gillberg, 2008; Howlin, Goode, Hutton, & Rutter, 2004), or burden of disease (Sanchez-Valle et al., 2008). These may reflect differences in their etiological factors, and physiopathologic and neuropathologic substrates (Rinehart et al., 2002).

Patients with HFA and AS clinically differ in language development (American Psychiatric Association, 1994). We might expect that volume abnormalities in classic language-processing regions in patients with HFA would be closer to those of patients with developmental language delay (DLD) or specific language impairment (SLI) and not present in those with AS (Bishop, 2010). Along these lines, one study reported similar volume abnormalities in patients with HFA and DLD compared with HC, such as a larger total GM volume, WM volume, and TBV (Herbert et al., 2005), with a positive correlation between IQ and TBV only in the DLD group, suggesting that brain enlargement does not mark an advantageous situation for ASD patients (Zeegers et al., 2009). Another study that compared patients with ASD with and without language impairment (ALI and ANLI, respectively) with patients with SLI and HC, reported that only the “language-impaired” groups (i.e., ALI and SLI groups) had decreased vermis and posterolateral cerebellar lobule volumes relative to ANLI and HC, which correlated with poorer language performance (Hodge et al., 2010). Additionally, three of the above-mentioned studies reporting qualitative differences between HFA and AS demonstrated variability between the two groups of patients in the anatomy of areas crucial to language development (Jou et al., 2010; McAlonan et al., 2008; Toal et al., 2010). Conversely, in their meta-analysis, Yu et al. did not find volume differences in classic language-processing regions between
AS and AD groups compared with HC, but volume patterns between the two categories were not directly compared (Yu et al., 2011).

4.3 Is it logical to assume a clear anatomical substrate for patients with ASD?

Effect sizes for sMRI volumetric and morphometric differences between patients and healthy subjects or between phenotypically different groups of patients are usually small, not only in the sMRI studies included in this systematic review, but in sMRI studies of psychiatric patients in general. This may explain the inconsistent and sometimes contradictory findings of this review. Yet, one wonders whether this is due to “poor sensitivity” of sMRI techniques, methodological and design limitations of sMRI studies, or both.

4.3.1 The “poor sensitivity” of sMRI techniques

In the year 2000, the Child Neurology Society and American Academy of Neurology stated that sMRI was “poorly sensitive” for routine developmental screening and screening specifically for autism (Filipek et al., 2000; Filipek et al., 1999). Furthermore, in the past two decades, many authors have concluded that sMRI could not be justified for diagnostic screening in these individuals, as many of them have not found any pathological brain changes or have reported only subtle or unremarkable abnormalities (Battaglia & Carey, 2006; Ekman et al., 1991; Garber & Ritvo, 1992; Shevell, Majnemer, Rosenbaum, & Abrahamowicz, 2001; Sokol & Edwards-Brown, 2004). This “poor sensitivity” could be due, firstly, to the fact that the use of manual or semiautomated methods, which focus on previously selected regions, preclude the exploration of other brain regions potentially involved in the etiology and clinical course of ASD; secondly, to the fact that traditional VBM analyses quantify changes in GM or WM volume between groups in a voxel-wise manner, using a univariate approach, such that each voxel is individually compared, thus losing information about possible multifaceted differences; and, thirdly, to the fact that complex neurobehavioral disorders such as ASD may be a result of a dysfunction in brain circuitry, secondary to brain dyssynchrony (Wickelgren, 2005) and to aberrant cortico-cortical and cortico-subcortical connectivity (Wass, 2011), rather than a result of volume abnormalities. However, this “poor sensitivity” may also be due to the fact that studies are conducted in clinically variable patient groups. Focal findings may be variable in different subjects (Herbert, 2004), so one wonders if it is logical to assume a clear anatomical substrate for such heterogenous patients (Polsek, Jagatic, Cepanec, Hof, &
Simic), and whether “poor sensitivity” of sMRI is a consequence of methodological and design limitations of this kind of studies.

4.3.2 Methodological and design limitations of sMRI studies

There are numerous methodology and design limitations of sMRI studies on ASD. Firstly, there is an absence of clearly defined hypotheses and objectives in many of the studies, as well as an absence of a clear definition of the AS and AD-HFA and LFA- clinical phenotype (Cohen, Volkmar, Anderson, & Klin, 1993). In this sense, studies use different diagnostic criteria, from DSM-III to ICD-10, combined or not with the administration of standardized diagnostic instruments such as the ADI-R (Autism Diagnostic Interview-Revised) and ADOS-G (Autism Diagnostic Observation Schedule-Generic) or different IQ cut-off points varying from 65 to 80 to define LFA (Frazier & Hardan, 2009). These different diagnostic criteria, previously defined by the authors, may condition the main findings and conclusions in each study. Secondly, many of these studies have small sample sizes (e.g. Casanova et al., 2009; Corbett et al., 2009; Freitag et al., 2009; Hardan, Libove, Keshavan, Melhem, & Minshew, 2009; Haznedar et al., 2006; Herbert et al., 2005; Hyde, Samson, Evans, & Mottron, 2010; Kates, Ikuta, & Burnette, 2009; Ke et al., 2009; Knaus et al., 2009; Mitchell et al., 2009; Rojas et al., 2004). Thirdly, as we have already mentioned, there is the effect of heterogeneity of samples in terms of age, sex, IQ, psychopharmacologic treatment, comorbidity with known genetic disorders, etc. For instance, differences in genetic constitution may be associated with differences in brain structure (Miles & Hillman, 2000). Concerning the age factor, studies vary from those including samples with children and adolescents (Chiu et al., 2008), to those including adult patients only (Dziobek, Bahnemann, Convit, & Heekeren, 2010), or a wide age spectrum from childhood to adulthood (Courchesne, Press, & Yeung-Courchesne, 1993). This is a relevant factor because ASD is considered a neurodevelopmental disorder and because of the strong influence of age on several brain structures (amygdala, cerebellum, TBV, etc.). Additionally, brain volume deviations have been shown quite consistently in ASD at certain ages (Hazlett et al., 2011; Schumann et al., 2010). Regarding the sex factor, many authors have studied male patients only (e.g., Hardan et al., 2009; Piven et al., 1995) subsamples of girls being very scarce and even excluded from statistical analyses. In fact, there is only one recent study that has investigated the volumetric differences between a sample entirely composed of female children with ASD and a sample of age- and IQ-matched HC, reporting greater GM volume in the left superior frontal gyrus in the ASD group (Calderoni et al., 2012). This is an important caveat, as several studies suggest that there are sex-specific differences in etiological factors and in the
The time course of ASD. For instance, one study reported that increased WM volume in frontoparietal, temporal, and cerebellar regions is more specific to females with ASD, in contrast with GM abnormalities, more commonly found in males (Craig et al., 2007). Another study described a “diagnosis x gender” effect for TBV, reporting larger volumes in autistic males but not females compared with HC (Piven, Arndt, Bailey, & Andreasen, 1996). Another one, comparing male and female patients with AS and HC reported that the typical sexual dimorphism found in HC, whereby males have larger total WM volume, was absent or attenuated in the AS group (Beacher et al., 2012). Finally, one study directly compared young boys and girls with autism, finding similar volume abnormalities in both groups but an age-structure volume relationship in girls compared with boys (Bloss & Courchesne, 2007). Studies using DTI also support sexual dimorphism in WM microstructure, which may account for the higher prevalence of ASD in males (Chou, Cheng, Chen, Lin, & Chu, 2011). Concerning IQ, it seems that it may be a confounding factor that may partly explain the different abnormality patterns between subgroups of patients with ASD (Salmond et al., 2007). Fourthly, many studies are cross-sectional, so they lose information about pathological developmental trajectories, particularly relevant in ASD. And, finally, sMRI studies use heterogeneous methodology in terms of measurement devices, manual tracing and data analysis techniques, and covariates used.

### 4.3.3 Potential solutions and the future of sMRI studies

The varied and recent sMRI applications (Stigler et al., 2011; Zeegers et al., 2006) and the combination of this technique with new neuroimaging approaches examining local organization and brain functional and structural connectivity may improve “sMRI sensitivity” (Caviness, Makris, Lange, Herbert, & Kennedy, 2000; Deeley & Murphy, 2009). This includes DTI techniques studying volume and micro-structural WM abnormalities in brain regions and connection tracts (Alexander et al., 2007; Barnea-Goraly et al., 2004; Mengotti et al., 2011), functional MRI focusing on abnormalities of functional connectivity and neuronal synchronization within and between neurofunctional regions and networks (Gepner & Feron, 2009), proton magnetic resonance spectroscopy studies, providing information regarding in vivo brain metabolite concentrations (Lauvin et al., 2012), magnetoencephalographic (MEG) studies, which can link neural activity to behavioral performance (Roberts et al., 2011), and electroencephalographic (EEG) registers, which can evaluate the efficiency of long- and short-range connections in ASD (Barttfeld et al., 2011). These combinations have already started to clarify the underlying neuroanatomical abnormalities
and brain-behavior relationships in ASD and other developmental disorders (Muller, 2007, 2008; Toal, Murphy, & Murphy, 2005).

Moreover, new classification methods based on machine learning methodology, are being developed in an attempt to overcome the limitations inherent in univariate VBM approaches, by making inferences about sets of brain voxels, that, in combination, can be used to discriminate between two participant groups (i.e., ASD or HC, HFA or AS, etc.) (Ecker, Marquand et al.; Ecker, Rocha-Rego et al.; Neeley et al., 2007; Uddin et al., 2011). Additionally, classification methods can investigate the relationship between symptom profiles and the key sets of brain voxels considered in the classification algorithm (Uddin et al., 2011). These new methods may help to develop diagnostic biomarkers by identifying those brain regions providing the greatest information regarding group membership.

On the other hand, further studies in well-defined and homogeneous patient populations sharing similar endophenotypes may better elucidate the pathophysiology and neurobiological underpinnings of ASD and its diagnostic subcategories (Frank & Pavlakis, 2001). Potential approaches to achieve this goal may include a) the development of a shared study scheme and a common ASD data system for different research groups, in order to provide comparable and reliable data; b) the design of prospective studies, with larger and more homogenous samples, including younger male and female naïve patients with ASD and well-matched healthy controls (i.e., gender-, age-, and IQ-matched); c) the supplementation of limited classification systems, such as the ICD-10 and DSM-IV, with gold-standard diagnostic assessments, such as the ADI-R and ADOS-G, that make it possible to score different psychopathological domains; and f) the enhancement of multidisciplinary research by combining neuroimaging data with clinical, neuropsychological, neuropathologic, neurochemical, and genetic data, not only in patients, but also in their relatives and healthy individuals (Anagnostou & Taylor, 2011; Davis et al., 2008; Hrdlicka et al., 2005; Levy et al., 2009; Palmen & van Engeland, 2004; Peterson et al., 2006; Posey & McDougle, 2001; Schumann & Nordahl, 2011; Tan, Doke, Ashburner, Wood, & Frackowiak, 2010; van Kooten et al., 2008; Wassink et al., 2007).

4.4 Conclusions

From the point of view of psychiatric taxonomy, Volkmar has already argued that the inclusion of a specific diagnosis, such as AS, within a nosological classification, “is only important if the use of the concept can be supported on the basis of some external validating factor” (Volkmar et al., 2000). For
example, specific clinical features have been used as arguments to justify maintaining AS in DSM-5 (Ghaziuddin, 2010), such as the presence of ego-dystonic lack of reciprocal social interaction in patients with AS, which may increase the risk for emotional comorbidities such as depression (Ghaziuddin, Weidmer-Mikhail, & Ghaziuddin, 1998; Strang et al., 2012), higher in AS than in AD. This argument may be considered at least a clinical specifier, as it may be relevant for prognosis and for treatment plan designing. However, it seems that brain structural data do not help resolve the controversy about whether or not to subsume DSM-IV categories under this new ASD diagnostic category proposed for DSM-5.

There are two reasons for this: firstly, because the scarcity, inconsistency, and methodological limitations of sMRI studies on ASD preclude findings from this type of study from providing a rationale for this proposal; and, secondly, because the value of sMRI data in autism research, if any, may depend on its associations with autism symptomatology and underlying neuropathology, but this has not yet been demonstrated sufficiently.

The fact that sMRI data do not help resolve this diagnostic controversy is not exclusive to ASD but applies to all psychiatric disease. For example, there is also a debate on whether autism and SLI are distinct disorders or should be considered similar categories within a “continuum,” as they share similar structural language impairments (i.e., autism is considered an “SLI plus,” assuming that the only factor differentiating the disorders is the presence of additional impairments in autism) (Bishop, 2003, personal communication). Here again, the lack of observed sMRI differences between SLI and ASD should play no role in determining whether or not the two disorders should be lumped together in a nosological classification. This controversy also occurs even with disorders from different nosological axes that are mutually exclusive but share similar diagnostic criteria, such as AS and Schizotypal Personality Disorder (SPD) (Esterberg, Trotman, Brasfield, Compton, & Walker, 2008; Hurst, Nelson-Gray, Mitchell, & Kwapis, 2007). Here again, there are still no sMRI data for determining whether or not they should be considered similar disorders.

In the end, the DSM-5 proposal has failed to be a neurobiological evidence-based classification system, as objective tools for assessing mental disorders have proven to be too rudimentary as compared with clinical methods for diagnosing patients and designing treatment plans. Therefore, future ASD research should be directed toward providing sensitive and specific biomarkers, which, in combination with behavioral tests, could easily lead to accurate and early diagnosis of the disorder. In order to achieve this, a plausible option would be to create different diagnostic classifications for research and for clinical
purposes, and, within the former, to define different research categories grouped by spectra, dimensions, endophenotypes, or other grouping criteria that may vary for different research questions (Tuchman, 2003), and that would be easier to change than clinical nosological classifications. Clinical classifications could then be modified when strong neurobiological evidence contradicts them.

**Conflict of interest statement**

The authors declare that there are no conflicts of interest.

**Acknowledgments**

Supported by the Spanish Ministry of Economy and Competitiveness, Instituto de Salud Carlos III, CIBERSAM, the Autonomous Community of Madrid, I+D Biomedicine, S2010/BMD-2422 AGES (Madrid, Spain), the ERA-NET NEURON (Network of European Funding for Neuroscience Research), Fundación Alicia Koplowitz and Fundación Mutua Madrileña.

**References**


