Brainstem enlargement in pre-school children with autism: results from an inter-method agreement study of segmentation algorithms.

Paolo Bosco[1], Alessia Giuliano[1], Jonathan Delafield-Butt[2], Filippo Muratori[3], Sara Calderoni[3] and Alessandra Retico[1]

[1] *INFN - National Institute for Nuclear Physics, Pisa Division, Pisa, Italy*

[2] *Faculty of Humanities and Social Science, University of Strathclyde, Glasgow, United Kingdom*

[3] *Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy; IRCCS Stella Maris Foundation, Pisa, Italy*

AUTHORS' ACCEPTED MANUSCRIPT

FOR PUBLICATION IN *HUMAN BRAIN MAPPING*

ACCEPTED FOR PUBLICATION 1ST AUGUST 2018

SEE https://onlinelibrary.wiley.com/journal/10970193 FOR THE FULL PUBLISHED PAPER

Corresponding author.

Paolo Bosco, paolo.bosco@pi.infn.it

Largo Bruno Pontecorvo 3, 56127 Pisa, Italy

Tel +39 050 2214397

Short title: Brainstem enlargement in children with autism

Keywords: Brainstem, Autism Spectrum Disorders, Segmentation, T1-weighted MRI, Brainstem volume.

Abstract

The inter-method agreement between automated algorithms for brainstem segmentation is investigated, focusing on the potential involvement of this structure in Autism Spectrum Disorders (ASD). Inconsistencies highlighted in previous studies on brainstem in the population with ASD may in part be a result of poor agreement in the extraction of structural features between different methods. A sample of 76 children with ASD and 76 age-, gender- and intelligence-matched controls was considered. Volumetric analyses were performed using common tools for brain structures segmentation, namely FSL-FIRST, FreeSurfer (FS), and Advanced Normalization Tools (ANTs). For shape analysis SPHARM-MAT was employed. Inter-method agreement was quantified in terms of Pearson correlations between pairs of volumes obtained by the different methods. The degree of overlap between segmented masks was quantified in terms of the Dice index. Both Pearson correlations and Dice indices, showed poor agreement between FSL-FIRST and the other methods (ANTs and FS), which by contrast, yielded Pearson correlations greater than 0.93 and average Dice indices greater than 0.76 when compared with each other. As with volume, shape analyses exhibited discrepancies between segmentation methods, with particular differences noted between FSL-FIRST and the others (ANT and FS), with under- and over-segmentation in specific brainstem regions. These data suggest that research on brain structure alterations should cross-validate findings across multiple methods. We reliably detected an enlargement of brainstem volume in the whole sample and in the male cohort across multiple segmentation methods, a feature particularly driven by the subgroup of children with idiopathic intellectual disability associated with ASD.

Introduction

Deficit in social communication abilities and the presence of restricted, repetitive behaviours represent the core features of autism spectrum disorder (ASD; American Psychiatric Association 2013). In addition, motor abnormalities have been consistently reported in ASD individuals (Fournier et al. 2010) as an early impairment (Teitelbaum et al. 1998) that may precede the development of defining characteristics (Ozonoff et al. 2008), is strongly correlated with the level of ASD severity (Hilton et al. 2007; Jasmin et al. 2009), and is associated with a poor prognosis (Sutera et al. 2007). However, the neural substrate underlying poor motor skills remains to be determined. Motor abilities depend on multiple interacting pathways including cortico-cortical, cortical-subcortical, and cortico-cerebellar connections that reach spinal motor neurons through the brainstem (Drew, Prentice, and Schepens 2004). From the anatomical point of view, the brainstem consists of midbrain (or mesencephalon), pons and medulla oblongata along the rostro-caudal axis, and is involved in several basic functions, including regulation of heart rate, breathing, alertness, sleeping, and eating (Angeles Fernández-Gil et al. 2010). It also plays a pivotal role in sensory information processing, in eliciting goal-oriented behaviour, in the regulation of social attention, and in the modulation of emotions (Berntson and Micco 1976; Geva et al. 2017; Venkatraman, Edlow, and Immordino-Yang 2017). Structural MRI studies of ASD individuals have demonstrated the potential role of the brainstem in the pathophysiology of ASD (Gaffney et al. 1988) (Roger J. Jou et al. 2013). Some of them have suggested a relationship between a reduction in brainstem volume and clinical abnormalities. In particular, a reduction in brainstem grey matter (GM) volume was reported in children with ASD compared to typically-developing (TD) controls, and a positive correlation between brainstem GM volume and oral sensory sensitivity was observed (R. J. Jou et al. 2009). Hanaie et al. (2016) detected a significantly smaller white matter (WM) volume in the brainstem of ASD children compared to TD controls which appears to be correlated with poor motor performance.

In a recent study, an inverse correlation between aggressive behaviour and brainstem volume was found in children with ASD (Lundwall et al. 2017). However, the findings in volume alterations in subjects with ASD with respect to matched controls are controversial, both in adults and child cohorts, with some early studies that did not detect any significant differences between the ASD and control samples (Elia et al. 2000; HARDAN et al. 2001; Herbert et al. 2003; Hsu et al. 1991; Kleiman, Neff, and Rosman 1992). More recently, a voxel-based morphometry (VBM) analysis revealed significantly increased GM and decreased WM in the brainstem of young adults with high-functioning ASD relative to TD controls (Hyde et al. 2010). A two-year longitudinal study investigating the brainstem volume showed an atypical development in children and adolescents with ASD compared to matched controls, characterized by an increase of GM brainstem volume over time (Roger J. Jou et al. 2013). Further, both quantitative and qualitative sex differences have been detected in the neuroanatomy of children (Retico et al., 2016) and adults (Ecker et al., 2017) with ASD, and the brainstem is part of such sexually dimorphic brain regions (Lai et al., 2013). Moreover, recent lesion studies suggested an involvement of the brainstem in neurocognition (D'Aes et al., 2015), and specifically a cognitive dysfunction after brainstem damage (Fu et al., 2017): thus, it is crucial to analyse the impact of IQ level on brainstem volume. None of the previous studies on brainstem structural MRI in ASD individuals has addressed this issue, and more broadly low-functioning children have been neglected in neuroanatomical autism research (Jack and Pelphrey, 2017). Moreover, the contribution of methodological variability to metrics of brainstem volume is still unclear.

In the present study, structural MRI was used to measure the volume and the shape of the brainstem in young children with ASD and matched controls. In addition, the impact of sex and intellectual functioning on brainstem morphometry was explored. Since the most recent studies elucidated an increase of brainstem GM volume in adolescents and young adults, we expected to detect some alterations also in young children. Moreover, since inter-method discrepancies can cause inconsistent

findings (Katuwal et al. 2016), we performed the segmentation of the brainstem structure with different methods, chosen among the most widespread analysis tools implemented in neuroimaging research studies. They include FIRST (Patenaude et al. 2011), FreeSurfer (Fischl et al. 2002; Iglesias et al. 2015) and ANTs (Advanced Normalization Tools; B. B. Avants et al. 2008). Moreover, we carried out the shape analysis of the segmented volumes, according to the SPHARM-MAT procedure (Brechbühler, Gerig, and Kübler 1995).

Material and methods

Dataset description

A group of 76 children with ASD, including 38 males [mean age ± SD = 53 ± 16 months; age range = 27-87 months] and 38 females [mean age ± SD = 53 ± 18; age range = 25-88 months] and a group of 76 control children matched by age, gender and non-verbal-IQ (NVIQ) were chosen for this case-control study. The same dataset has been previously analysed to investigate sex-related structural differences in young children with ASD (Retico et al. 2016) and the inclusion and exclusion criteria for subjects with ASD and controls were exhaustively described beforehand (Retico et al. 2016; Calderoni et al. 2012). Participants in the ASD group were recruited at the Autism Spectrum Disorders Unit of IRCCS Stella Maris Foundation (Pisa, Italy), a tertiary care university hospital. They met the criteria for diagnosis in the autism spectrum according to DSM-5, and underwent a MRI scan. The control group was constituted by 38 children (19 males and 19 females) with intellectual disability (ID), i.e. with NVIQ score <70, and 38 children without ID (no-ID, 19 males and 19 females), i.e. with NVIQ score ≥70. Children with ID were included within the control group to guarantee the match for NVIQ between children with ASD and controls. The control group of subjects with ID was very accurately selected, as described elsewhere (Retico et al. 2016).

The dataset composition has been summarized in Table 1, where the mean, the standard deviation and the range of age and NVIQ of each group have been reported. This study was approved by the Institutional Review Board of the Clinical Research Institute for Child and Adolescent Neurology and Psychiatry, and the written informed consent was obtained from all parents or tutors.

Clinical procedures and measures

Children with ASD underwent a multidisciplinary team evaluation, including a senior child psychiatrist, a child psychologist, a speech-language pathologist and an educational therapist, during 5-7 days of extensive assessment. In particular, all children with ASD performed the following assessments.

A number of well-standardized tests were used to assess intellectual abilities due to differences in the age, verbal skills and functioning level of children. These evaluations included: the Leiter International Performance Scale-Revised (Roid and Miller 1997), the Griffiths Mental Development Scale (Griffiths 1984), the Italian version of Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, Scales, and Index 2012) and Wechsler Intelligence Scales for Children-Revised (WISC-R; Wechsler 1974). When the tool provides a mental age (MA), NVIQ was estimated dividing MA by the child's chronological age (CA): ([MA/CA] x 100). In this study, we consider the NVIQ scores.

The Autism Diagnostic Observation Schedule-Generic (ADOS-G; Lord et al. 2000) was employed to evaluate social and communicative functioning in children suspected of having an ASD. Module 1 of the ADOS (designed for preverbal children or children with only single words) or Module 2 (children with consistent phrase speech) was administered by two clinical psychologists who met standard requirements for research reliability. The average value, the standard deviation and the

range of the ADOS-G have been reported in Table 1 for the entire sample of subjects affected by ASD and for the male and female subsamples separately.

MRI data acquisition

All MRI scans were acquired in the same tertiary care hospital using a 1.5 T MR Neuro-optimized System (GE HealthCare, USA) fitted with 40 mT/m high-speed gradients. The standard MR protocol for children included a whole brain Fast Spoiled Gradient Recalled acquisition in the steady-state T1-weighted series (FSPGR). Isotropic images were collected in the axial plane with repetition time 12.4 ms, echo time 2.4 ms, inversion time 700 ms, flip angle of 10, yielding 124 contiguous 1.1 mm slices with an in-plane resolution of 1.1x1.1 mm^2. Both ASD and control children were sedated with a general anaesthesia with a halogenated agent while spontaneously breathing, according to a strict clinical protocol approved by the institutional review board of the IRCCS Stella Maris Foundation and performed by a paediatric anaesthesiologist.

Automated methods for brainstem segmentation

To segment the brainstem we used FIRST (Patenaude et al. 2011; FSL version 5.0.8), FreeSurfer (Fischl et al. 2002) (version 5.3, 6.0 and 6.0 with brainstem substructure extraction) and ANTs (Advanced Normalization Tools; B. B. Avants et al. 2008).

Brainstem segmentation with FIRST was performed through the standard run_first_all command. FIRST consists in a probabilistic atlas approach which uses manually labelled image data to provide anatomical training information. It utilises the principles of the Active Shape and Appearance Models (i.e. images are segmented using the model built from the training data, which specifies the range of likely shapes) but places them within a Bayesian framework, allowing probabilistic relationships between shape and intensity to be fully exploited. The model is trained for 15 different subcortical structures using 336 manually-labelled T1-weighted MR images. Prior to building the shape and intensity model, all data is registered to a common space based on the non-linear MNI152

template with 1 × 1 × 1 mm resolution. It is important to remark that FIRST brainstem model includes the fourth ventricle. Indeed, the fact that the fourth ventricles passes directly through the brainstem makes the shape approach more complex and challenging.

The pipeline recon-all was run for FreeSurfer 5.3 and FreeSurfer 6.0. The first step in the FreeSurfer processing is to perform a motion correction, affine transformation to Talairach image space, non-uniform intensity normalization for intensity inhomogeneity correction, and removal of non-brain tissues. The remaining brain image volume is intensity normalized to match the FreeSurfer atlas image intensity histogram, which is followed by a non-linear warping of the atlas brain image to subject brain image. Warped atlas brain image in the subject image space is utilized in atlas-based tissue segmentation, in labelling the subcortical structures, in our case the brainstem structure. Except for bug fixing, the developers did not declare significant differences between FS.5.3 and FS6.0 for subcortical structures segmentation, therefore in principle no significant differences were expected.

The additional pipeline for the brainstem substructure segmentation (Iglesias et al. 2015), available for FreeSurfer 6.0 only, was run to obtain the volumes of both the total brainstem and its substructures: the medulla oblongata (MO), the pons, the midbrain and the superior cerebellar peduncle. This segmentation method relies on a probabilistic atlas of the brainstem and its neighbouring brain structures. The atlas consists in a dataset of 49 manually labelled with a protocol that was specifically designed for the study and fully described in (Iglesias et al. 2015). Importantly, this protocol includes the superior cerebellar peduncles, thus significant differences in brainstem segmentations were expected with respect to the other methods.

A specific pipeline analysis, based on multi-atlas segmentation approach, was developed by using the utilities provided by ANTs. First of all, a dataset-specific sMRI template was created (antsMultivariateTemplateConstruction.sh script (Brian B. Avants et al. 2011)). Then the whole data

set was registered to the template with a symmetric diffeomorphic transform (antsRegistrationSyNQuick.sh script (B. B. Avants et al. 2008)). Then, we used the multi-atlas label fusion (MALF) algorithm (Wang et al. 2013) in conjunction with a manually labelled (Klein and Tourville 2012) subset of 20 OASIS (Marcus et al. 2007) scans, to generate the brainstem segmentation. First, we normalized the labelled OASIS subset to the template. Then, we performed MALF on each subject using the normalized labelled data as input, thus obtaining the brainstem segmentation in the template space. An inverse transform was finally applied to take the brainstem labels back to the native space.

In order to perform direct comparisons on both brainstem volumes and shapes, every segmentation was analysed in the native space.

Alongside brainstem segmentations, the total intracranial volume of each subject was extracted by SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/). We chose SPM12 since it was independent from all the considered brainstem segmentation methods and it already proved to be one of the most reliable and accurate tools for total intracranial volume (TIV) estimation (Malone et al. 2015).

Statistical analysis

A direct comparison of brainstem volumes in native space was performed to evaluate inter-method agreement (correlation was calculated by Pearson coefficient) and Dice similarity indexes (range [0,1], where 1 means total agreement) (Dice 1945) were calculated to evaluate the segmentation similarity across methods.

The statistical examination of brainstem volume was performed using the analysis of variance (ANOVA) test. A comparison between ASD and control subjects was performed both in the male and female groups separately and in the entire dataset using gender as covariate. In addition, the impact of intellectual functioning on brainstem volumes was explored. In all cases TIV and age were

introduced as covariates in the statistical model to take into account both the inter-individual head size variability and the age dependence of the brain structures. A *p*-value of 0.05 was considered as significance level to reject the null hypothesis of equal means between the groups. ANOVA tests were also performed to assess volume differences between ASD and controls at the level of brainstem substructure volumes estimated using FreeSurfer 6.0. Since these volumes represent a set of correlated variables, the significance level was Bonferroni corrected, and a *p*-value of 0.01 was considered as significance threshold.

We investigated the correlation between the brainstem volume, normalized with respect to TIV, and the ASD symptom severity ADOS-G total scores. The Pearson's correlation index was estimated to investigate the evidence against the null hypothesis of no correlation.

Shape analysis

We performed a shape analysis of the brainstem segmentations by using a standard SPHARM-MAT procedure (Brechbühler, Gerig, and Kübler 1995), which creates parametric surface models using spherical harmonics. The workflow was composed of the following steps: a) a pre-processing step of the 3D binary data which assures that the voxel surface has a spherical topology; b) a spherical parameterization, i.e. the calculation of a continuous and uniform mapping from the object surface to the surface of a unit sphere; c) a spherical parametrization expansion by which the object surface description was expanded into a complete set of spherical harmonic basis functions; d) Surface alignment which registered the 3D spherical models in a common space in a two steps process. First, a first order ellipsoid (FOE) (Styner et al. 2006) alignment method was applied. Second, the registration was refined by SHREC algorithm (Shen, Farid, and McPeek 2009). In this step, we used the mean surface of the FOE aligned surfaces as spatial reference. After registering all the individual SPHARM models to the template model, all the SPHARM coefficients are comparable across objects, and then group analysis can be performed. In particular, each SPHARM surface was

uniformly sampled to create a landmark representation. Then, the deformation along normal direction of the atlas surface was calculated for each landmark, providing for each subject a measure of its local similarity to the reference surface (Shen et al. 2006).

For statistical models of the signal extracted on the surfaces, we used SurfStat (Worsley et al. 2009), a free software tool which performs statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effect models and random field theory.

Results

Brainstem segmentation agreement across methods

The 3D rendering of brainstem segmentations obtained with the 5 different methods (ANTs, FSL-FIRST, FreeSurfer 5.3, 6.0, and 6.0 with substructures, respectively) for a sample subject are shown in figure 1. The renderings are overlaid onto the corresponding median sagittal view of the T1-weighted MRI.

The comparisons between methods in terms of scatter plots of brainstem volumes provided by the algorithms under investigations are reported in figure 2. The colours represent the Dice coefficients computed for each pair of segmented masks in native space. Alongside the scatter plots, the Pearson correlation coefficients (R) are reported to summarize the overall agreement in terms of volumetric information. Moreover, the mean Dice indexes (D) and their Standard Deviations are indicated to summarize the agreement between different methods in terms of overlap of the brainstem segmentations.

Dice similarity index ranges from 0.53 ± 0.14 (FSL-FIRST vs FS 6.0 with substructures) to 0.93 ± 0.07 (FS 5.3 vs FS 6.0). In particular, every comparison involving FSL-FIRST shows a poorer agreement with all the other methods, with an estimate of the volume that is lower with respect to the

other algorithms. Consistently, the average Pearson correlation coefficient is greater than 0.9 in all cases except for those involving FSL-FIRST (ranging from 0.5 – 0.65).

The corresponding shape analyses by means of Student's T maps of the deformation along normal direction from a reference surface are shown in figure 3. As reference surface, the mean surface of the whole data set segmented by the couple of methods under investigation was considered. Warmer colours denote a positive distance from the mean surface, i.e. on average the area is over-segmented by the first methods with respect to the second. Conversely, cooler colours denote a negative distance from the reference surface, i.e. an under-segmentation of the first algorithm with respect to the second one.

T maps were thresholded at cluster-level p-value <0.05 corrected according to Random-Field Theory (RFT).

Shape analysis confirms that different segmentation methods significantly differ from each other. The differences (see figure 3) are particularly striking with FSL-FIRST method, where the estimate of the volume is in general lower, except for the systematic higher estimate of the brainstem structure in the posterior part corresponding to the fourth ventricle which is included in FSL-FIRST brainstem segmentation protocol. Regarding ANTS (first row in figure 3), the pons area looks systematically over-segmented with respect to the other methods (warmer colours from yellow to red located frontally in the brainstem structures). As expected, the most similar brainstem segmentations are those obtained by FS 5.3 and FS 6.0 (in the third row of figure 3, the first two T-maps do not show significant differences), while the FS 6.0 substructure segmentation (last column in figure 3) shows significant differences, due to the inclusion of the superior cerebellar peduncle structures that are not included in the segmentation protocols of the other methods. On the contrary, the comparisons of the anterior areas suggest a systematic reduction in pons volume segmentation with respect to ANTs, FS 5.3 and FS 6.0 methods.

Statistical analysis of brainstem volumes

The results of comparison between ASD and control subjects of brainstem total volumes computed using ANTs, FSL-FIRST, FS 5.3, FS 6.0 and the sum of FS 6.0 substructures, are summarized in Table 2. TIV, age and gender were considered as covariates (significance level $p<0.05$) and gender-specific subgroup differences were also assessed. The brainstem volume resulted significantly higher in ASD when compared to controls both in the entire sample and in male subgroup for ANTs, FS 5.3, and FS 6.0 computations. Instead, no statistically significant volume differences were obtained for females in brainstem total volume. ANOVA test was also performed to look for brainstem substructure volume differences between ASD and controls. There were no statistically significant differences when accounting for TIV, age and gender and applying Bonferroni correction (significance level $p<0.01$) between ASD and control subjects.

The brainstem total volume and its substructures were statistically compared between ASD individuals and controls in the subgroups of subjects with ID (NVIQ<70) and without ID (NVIQ≥70), each of them composed by 76 subjects. The tests were performed both in the entire sample and in the gender-related subgroups, with the same covariates as above. As evident from Table 3, brainstem volume and its sub-regions tend to be larger in ASD when compared with controls both in ID and in no-ID groups. However, there were no statistically significant differences in the group of subjects without ID both in brainstem total volume and in substructures. On the contrary, when considering only subjects with NVIQ<70 a number of significant differences were detected. The total brainstem resulted to be significantly higher in ASD with respect to controls ($p<0.05$) in the entire sample (for ANTs, FS 5.3 and FS 6.0), and in male subgroup for FS 5.3 and FS 6.0. Moreover, the superior cerebellar peduncle was significantly higher in ASD than controls ($p<0.01$ with Bonferroni correction). Instead, in female subgroup with ID there were no statistically significant differences both in total brainstem volume and in substructures.

In all considered comparisons (both in Table 2 and Table 3), no significant volume differences were identified when brainstem segmentation was performed by FSL-FIRST algorithm.

The study of the correlation between the brainstem volumes normalized to TIV and the ADOS total scores was carried out for the male, the female and the whole sample of ASD children. Five female children were excluded from this test because of the absence of their ADOS-G score. No significant results were obtained for any of the tests against the null hypothesis of no correlation on $p<0.05$ for the entire dataset analysis and for the gender-related subgroups.

Shape analysis

Given that some volume differences were detected through statistical comparisons by means of ANOVA, we performed a shape analysis to detect brainstem shape differences between ASD and CTRL groups in the very same way used to compare the segmentations obtained by the different methods. Since ANTs and FreeSurfer-based methods showed a good agreement in terms of volumetric information, we chose to show the shape analysis on the brainstem segmentations obtained by FS 6.0 with substructures. Indeed, it is the most recent method with the most inclusive segmentation protocol among the others. In Appendix A the very same shape analyses are reported for both FSL-FIRST and ANTS methods.

For each group comparison, we calculated a linear mixed effects model to describe the distance along normal direction from the reference surface, calculated for each vertex of the mesh describing the surface. In analogy to the ANOVA analysis, we included in the linear model the TIV, the age and the gender of each subject. We used as reference surface the average surface of the brainstem segmentation of the CTRL cohort. We imposed a threshold on the T maps based on cluster-level p-value <0.05 RFT corrected.

The comparison between the shapes of the brainstem segmentations for ASD and CTRL subjects is shown in figure 4a. Statistically significant differences can be detected. In particular, the T map shows an overall inflation of the brainstem surfaces of ASD cohort with respect to the reference surface. The most involved areas are in the posterior side of the brainstem involving the superior cerebellar peduncle and medulla oblongata substructures.

Figures 4b and 4c show the shape comparison between ASD and CTRL groups for the ID and no-ID subgroups, respectively. Consistently with the ANOVA results, statistically significant brainstem surface inflation can be detected in the subgroup of the ASD subjects with ID with respect to the matched CTRL group. Analogously to the previous case, the most involved areas are located in the posterior side of the brainstem, involving the medulla oblongata. On the contrary, no statistically significant differences in terms of brainstem shapes were detected in the no-ID subgroup.

Discussion

To tackle inconsistent findings on subcortical structure alterations in ASD, we performed the segmentation of the brainstem structure with different automated methods, chosen among the most widespread analysis tools (ANTs, FSL-FIRST, FreeSurfer-based methods), and carried out shape analysis of the segmented volumes with SPHARM-MAT.

Both in terms of the Pearson correlation and of the Dice index, FSL-FIRST showed poor agreement with the other segmentation methods (ANTs and FS-based methods), which, by contrast, consistently yielded Pearson correlations greater than 0.93 and average Dice indices greater than 0.76 in comparisons between each other. Inter-method brainstem volume differences can be attributed to varying definitions of brainstem structure, the use of different templates (e.g. in our study only ANTs processed the brain scans by using an age-specific brain template), and the varying effects of imaging artefacts and acquisition settings.

The shape analysis confirms the discrepancies among different segmentation methods, with particular reference to the FSL-FIRST under- and over-segmentation problems in specific brainstem regions. In particular, an over-segmentation of the brainstem close to the 4th ventricle can be identified with respect to all the other considered methods. The overall lower estimation of the volume could both be due to the FSL_FIRST segmentation method and to a different manual protocol of the labelled images used in FSL-FIRST training data. Indeed, as shown in Figure 3, the FSL-FIRST segmentation is tighter with respect to the other methods in almost the whole surface, with the exception of the 4th ventricle.

The volume analyses suggest that brainstem structure in the age range 22-89 months is larger in subjects with ASD with respect to controls. The alteration of brainstem volume is statistically significant when considering male and female populations together or the male population alone. It is possible that the brainstem of young children with ASD follows an early atypical development characterized by an increased volume in the first years of life, similarly to the precocious enlargement of the whole brain (Courchesne et al. 2007), and of distinct brain regions (Bellani et al. 2013; Carper et al. 2002; Carper and Courchesne 2005; Giuliano et al. 2017; Qiu et al. 2016) detected in individuals with ASD compared to matched controls. However, a longitudinal study rather than a cross-sectional one should be planned in young subjects with ASD to identify the brainstem growth trajectory of this population.

By contrast, despite that ANOVA estimated mean volumes were always slightly larger for ASD female children with respect to controls, this brainstem difference did not reach a statistically significant threshold in the female population alone. Sexual dimorphism in the neuroanatomy of children with ASD has been previously detected in several brain regions, including the corpus callosum (Qiu et al. 2016), the amygdala (Schumann et al., 2009), the temporal regions and the cerebellum (Bloss and Courchesne 2007), and the bilateral frontal regions (Retico et al. 2016).

Crucially, to our knowledge, no previous study has investigated the impact of sex on the brainstem volume of young individuals with ASD.

The shape analysis on FreeSurfer 6.0 with substructures segmentation confirms the enlargement of the brainstem suggesting that the main differences are located in the posterior side, involving the medulla oblongata and the superior cerebellar peduncles structures. These appear to include the posterior midline nuclei of the vagus, intercalatus, and hypoglossal nerves responsible for autonomic regulation of the thoracic organs, eye-gaze, and movement of the hypoglossus during speech. These physiological features are widely reported disrupted in ASD. Eye-gaze in particular has received significant attention in the literature, although cortex is often referred to as the site of neural disruption (Grice et al. 2005; Pelphrey, Morris, and McCarthy 2005; Jones and Klin 2013), rather than brainstem. This medullary region also holds the nucleus ambiguous concerned with regulation of vagal function in conjunction with the so-called social engagement system that makes up autonomic co-regulation through the 'polyvagal system' (Porges 2011), a discreet evolutionary adaptation in humans that allows for expressive information in gesture and speech to be coupled to autonomic arousal, giving its vital expressive qualities from infancy onward (Porges and Furman 2011). These features are reported disrupted in ASD individuals (Rochat et al. 2013).

The inferior olivary nucleus that sits posterior to the cerebrospinal fasciculus at the anterior side is adjacent to sites of enlargement reported here and is reported to be disrupted in ASD from neuromotor recordings and in post-mortem histology (Welsh, Ahn, and Placantonakis 2005; Fatemi et al. 2012), and may be contribute to the motor deficits reported in ASD (Fournier et al. 2010; Teitelbaum et al. 1998; Ozonoff et al. 2008). Similarly, the posterior superior cerebellar peduncles found enlarged here contain a number of tracts also responsible for sensorimotor information, namely the cerebello-thalamic tract responsible for proprioceptive information and communication with thalamus and cortex for voluntary control, and the cerebello-rubral and cerebello-reticular tracts that carry sensorimotor information between cerebellum and brainstem motor nuclei. Enlargement of

these regions stands in agreement with the notion that sensory and motor information integration is disrupted at the level of the brainstem, but the specific nature of this disruption is yet to be elucidated (Fatemi et al. 2012; Trevarthen and Delafield-Butt 2013).

It is important to note brainstem integration of information is not a passive event. New theoretical perspective suggests these basic neural systems give rise to a primary, core psychological experience of Self, the basic experience of being a person with feelings, intentions, and agency to act in the world with one's own purpose (Merker 2007; Panksepp and Biven 2012; Vandekerckhove and Panksepp 2011). The data we present in this paper begin to elucidate with greater clarity the brainstem system involved, adding resolution to an understanding of autism as a disruption to basic experience conveyed in the form of movement for communication and function, and in adaptive autonomic regulation of the body (Trevarthen and Delafield-Butt 2013; Delafield-Butt and Trevarthen 2017).

Interestingly, volume analyses show that brainstem enlargement is preeminent in ASD-ID subjects with respect to NVIQ-matched controls. This finding is consistent across ANTs and FreeSurfer-based segmentation methods. Analogous to the previous case, the effect is stronger in the male cohort and it remains statistically significant also when considering both male and female populations together. In contrast, there are no statistically significant differences between brainstem volumes of ASD-no-ID and NVIQ-matched control children (both in male and in female populations). The same result was obtained with shape analysis. In particular, FS 6.0 segmentations showed an inflation of ASD brainstem structure in ASD-ID subjects with respect to NVIQ-matched controls, particularly in medulla oblongata and superior cerebellar peduncle areas. No statistically significant alterations were detected between ASD-no-ID and no-ID controls. This result is in line with a recent study in which larger volumes in the ASD group were related to poorer performance (Lalani et al., 2018). Further, this finding is interesting in that recent investigations suggested a participation of the brainstem in cognitive processes (D'Aes et al., 2015; Fu et al., 2017), and points

at the relevance of further investigating the relationship between brainstem volume and intellectual functioning in children with ASD.

Among the limitations of the present study is the use of 1.5 T T1-weighted MRI, which shows a suboptimal image resolution and it is not fully appropriate to highlight a contrast between brainstem substructures. Other magnetic resonance acquisition sequences and their combination have demonstrated enhanced sensitivity to depict brainstem anatomy *in vivo*. For example, a study carried out by Lambert et al. at 3 T with 0.8 mm isotropic resolution exploited the magnetization transfer and proton density acquisitions to obtain an accurate multichannel segmentation of the brainstem and its subregions (Lambert et al. 2013). In addition to the possible improvement of brainstem segmentation accuracy, the use of other MRI-based acquisition modalities, e.g. diffusion imaging, has already highlighted complementary information about brainstem structural abnormalities in ASD. The investigation by Travers et al. (Travers et al. 2015) suggests that brainstem white matter contributes to autism symptoms and motor skills in ASD. In addition to the investigation of brainstem structure, functional studies based on resting-state functional connectivity (FC) approaches highlighted an increased FC between the putamen and a brainstem region in the pons in children with ASD (Di Martino et al. 2011). Finally, a major improvement to structural and functional studies of the brainstem could be provided by ultrahigh-field MRI (7 T and higher), thanks to its improved ability in depicting the ultrastructure of tissues and to its enhanced BOLD signal (Sclocco et al. 2017).

Conclusions

The reproducibility across different studies of possible alterations in the neuroanatomy of subjects with ASD may be strongly affected by the agreement of automated segmentation algorithms. The comparison among the segmentation outputs of FSL-FIRST, FreeSurfer-based methods and ANTs

showed Pearson correlations greater than 0.93 and average Dice indexes greater than 0.76, except for comparisons involving FSL-FIRST. The shape analysis based on SPHARM-MAT confirmed the same trend. The brainstem volume resulted significantly higher in children with ASD when compared to controls both in the entire sample and in male subgroup, where the result was driven by the subgroup of subjects with intellectual disability.

Acknowledgments

This work has been partially supported by the Tuscany Government (PAR-FAS 2007-2013, Bando FAS Salute 2014) through the ARIANNA Project (C52I16000020002) and by grant from the IRCCS Stella Maris Foundation (Ricerca Corrente, and the 5 × 1000 voluntary contributions, Italian Ministry of Health).

References

American Psychiatric Association, Task Force on DSM-IV. 2013. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV*. https://books.google.it/books?hl=it&lr=&id=-JivBAAAQBAJ&oi=fnd&pg=PT18&ots=ceNK-1MFBa&sig=21pJJ2I3NcLY67qKFv4f-W-pmhw#v=onepage&q&f=false (July 14, 2017).

Angeles Fernández-Gil, M., R. Palacios-Bote, M. Leo-Barahona, and J. P. Mora-Encinas. 2010. "Anatomy of the Brainstem: A Gaze into the Stem of Life." *Seminars in Ultrasound, CT and MRI* 31(3): 196–219. http://www.ncbi.nlm.nih.gov/pubmed/20483389 (September 19, 2017).

Avants, B. B., C. L. Epstein, M. Grossman, and J. C. Gee. 2008. "Symmetric Diffeomorphic Image Registration with Cross-Correlation: Evaluating Automated Labeling of Elderly and Neurodegenerative Brain." *Medical Image Analysis* 12(1): 26–41. http://www.ncbi.nlm.nih.gov/pubmed/17659998 (April 4, 2017).

Avants, Brian B. et al. 2011. "A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration." *NeuroImage* 54(3): 2033–44. http://www.ncbi.nlm.nih.gov/pubmed/20851191 (April 4, 2017).

Bellani, M, S Calderoni, F Muratori, and P Brambilla. 2013. "Brain Anatomy of Autism Spectrum Disorders II. Focus on Amygdala." *Epidemiology and psychiatric sciences* 22(4): 309–12. http://www.ncbi.nlm.nih.gov/pubmed/23815810 (September 19, 2017).

Berntson, Gary G, and David J Micco. 1976. "Organization of Brainstem Behavioral Systems." *Brain Research Bulletin* 1(5): 471–83. http://www.ncbi.nlm.nih.gov/pubmed/1034494 (September 19, 2017).

Bloss, Cinnamon S., and Eric Courchesne. 2007. "MRI Neuroanatomy in Young Girls With Autism." *Journal of the American Academy of Child & Adolescent Psychiatry* 46(4): 515–23.

http://www.ncbi.nlm.nih.gov/pubmed/17420687 (September 19, 2017).

Brechbühler, Ch., G. Gerig, and O. Kübler. 1995. "Parametrization of Closed Surfaces for 3-D Shape Description." *Computer Vision and Image Understanding* 61(2): 154–70. http://linkinghub.elsevier.com/retrieve/pii/S1077314285710132 (April 21, 2017).

Calderoni, Sara et al. 2012. "Female Children with Autism Spectrum Disorder: An Insight from Mass-Univariate and Pattern Classification Analyses." *NeuroImage* 59(2): 1013–22.

Carper, Ruth A., and Eric Courchesne. 2005. "Localized Enlargement of the Frontal Cortex in Early Autism." *Biological Psychiatry* 57(2): 126–33. http://www.ncbi.nlm.nih.gov/pubmed/15652870 (September 19, 2017).

Carper, Ruth A, Pamela Moses, Zachary D Tigue, and Eric Courchesne. 2002. "Cerebral Lobes in Autism: Early Hyperplasia and Abnormal Age Effects." *NeuroImage* 16(4): 1038–51. http://www.ncbi.nlm.nih.gov/pubmed/12202091 (September 19, 2017).

Courchesne, Eric et al. 2007. "Mapping Early Brain Development in Autism." *Neuron* 56(2): 399–413. http://www.ncbi.nlm.nih.gov/pubmed/17964254 (September 19, 2017).

D'Aes, T., and Marien, P. (2015). Cognitive and affective disturbances following focal brainstem lesions: a review and report of three cases. Cerebellum 14, 317–340. doi: 10.1007/s12311-014-0626-8

Delafield-Butt, Jonathan, and Colwyn Trevarthen. 2017. "On the Brainstem Origin of Autism: Disruption to Movements of the Primary Self." In *Autism: The Movement Sensing Perspective*, ed. Taylor & Francis CRC Press.

Dice, Lee R. 1945. "Measures of the Amount of Ecologic Association Between Species." *Ecology* 26(3): 297–302. http://doi.wiley.com/10.2307/1932409 (April 4, 2017).

Drew, Trevor, Stephen Prentice, and Bénédicte Schepens. 2004. "Cortical and Brainstem Control of

Locomotion.” In , 251–61. http://linkinghub.elsevier.com/retrieve/pii/S0079612303430252 (July 14, 2017).

Elia, Maurizio et al. 2000. “Clinical Correlates of Brain Morphometric Features of Subjects With Low-Functioning Autistic Disorder.” *Journal of Child Neurology* 15(8): 504–8. http://www.ncbi.nlm.nih.gov/pubmed/10961787 (July 14, 2017).

Fatemi, S. Hossein et al. 2012. “Consensus Paper: Pathological Role of the Cerebellum in Autism.” *The Cerebellum* 11(3): 777–807. http://link.springer.com/10.1007/s12311-012-0355-9 (November 27, 2017).

Fischl, Bruce et al. 2002. “Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain.” *Neuron* 33(3): 341–55. http://www.ncbi.nlm.nih.gov/pubmed/11832223.

Fournier, Kimberly A. et al. 2010. “Motor Coordination in Autism Spectrum Disorders: A Synthesis and Meta-Analysis.” *Journal of Autism and Developmental Disorders* 40(10): 1227–40. http://www.ncbi.nlm.nih.gov/pubmed/20195737 (July 14, 2017).

Fu X, Lu Z, Wang Y, Huang L, Wang X, Zhang H, Xiao Z. A Clinical Research Study of Cognitive Dysfunction and Affective Impairment after Isolated Brainstem Stroke. Front Aging Neurosci. 2017 Dec 19;9:400.

Gaffney, Gary R, Samuel Kuperman, Luke Y Tsai, and Susan Minchin. 1988. “Morphological Evidence for Brainstem Involvement in Infantile Autism.” *Biological Psychiatry* 24(5): 578–86. http://www.ncbi.nlm.nih.gov/pubmed/3167146 (June 21, 2017).

Geva, Ronny et al. 2017. “Brainstem as a Developmental Gateway to Social Attention.” *Journal of Child Psychology and Psychiatry and Allied Disciplines*. http://www.ncbi.nlm.nih.gov/pubmed/28504308 (September 19, 2017).

Giuliano, Alessia et al. 2017. “The Effect of Age, Sex and Clinical Features on the Volume of Corpus Callosum in Pre-Schoolers with Autism Spectrum Disorder: A Case-Control Study” ed. John Foxe. *European Journal of Neuroscience*. http://www.ncbi.nlm.nih.gov/pubmed/28112456 (September 19, 2017).

Grice, Sarah J. et al. 2005. “Neural Correlates of Eye-Gaze Detection in Young Children with Autism.” *Cortex* 41(3): 342–53. http://www.sciencedirect.com/science/article/pii/S0010945208702715?via%3Dihub (November 27, 2017).

Griffiths, R. 1984. “The Abilities of Young Children: A Comprehensive System of Mental Measurement for the First Eight Years of Life.”

Hanaie, Ryuzo et al. 2016. “White Matter Volume in the Brainstem and Inferior Parietal Lobule Is Related to Motor Performance in Children with Autism Spectrum Disorder: A Voxel-Based Morphometry Study.” *Autism Research* 9(9): 981–92. http://www.ncbi.nlm.nih.gov/pubmed/26808675 (July 14, 2017).

HARDAN, ANTONIO Y., NANCY J. MINSHEW, KEITH HARENSKI, and MATCHERI S. KESHAVAN. 2001. “Posterior Fossa Magnetic Resonance Imaging in Autism.” *Journal of the American Academy of Child & Adolescent Psychiatry* 40(6): 666–72. http://linkinghub.elsevier.com/retrieve/pii/S0890856709604706 (June 21, 2017).

Herbert, M R et al. 2003. “Dissociations of Cerebral Cortex, Subcortical and Cerebral White Matter Volumes in Autistic Boys.” *Brain : a journal of neurology* 126(Pt 5): 1182–92. http://www.ncbi.nlm.nih.gov/pubmed/12690057 (July 14, 2017).

Hilton, Claudia et al. 2007. “Relationship between Motor Skill Impairment and Severity in Children with Asperger Syndrome.” *Research in Autism Spectrum Disorders* 1(4): 339–49. http://linkinghub.elsevier.com/retrieve/pii/S1750946707000037 (July 14, 2017).

Hsu, M, R Yeung-Courchesne, E Courchesne, and G A Press. 1991. "Absence of Magnetic Resonance Imaging Evidence of Pontine Abnormality in Infantile Autism." *Archives of neurology* 48(11): 1160–63. http://www.ncbi.nlm.nih.gov/pubmed/1953402 (July 14, 2017).

Hyde, Krista L., Fabienne Samson, Alan C. Evans, and Laurent Mottron. 2010. "Neuroanatomical Differences in Brain Areas Implicated in Perceptual and Other Core Features of Autism Revealed by Cortical Thickness Analysis and Voxel-Based Morphometry." *Human Brain Mapping* 31(4): 556–66. http://www.ncbi.nlm.nih.gov/pubmed/19790171 (July 14, 2017).

Iglesias, Juan Eugenio et al. 2015. "Bayesian Segmentation of Brainstem Structures in MRI." *NeuroImage* 113: 184–95. http://www.ncbi.nlm.nih.gov/pubmed/25776214 (April 4, 2017).

Jack A, A Pelphrey K. Annual Research Review: Understudied populations within the autism spectrum - current trends and future directions in neuroimaging research. J Child Psychol Psychiatry. 2017 Apr;58(4):411-435.

Jasmin, Emmanuelle et al. 2009. "Sensori-Motor and Daily Living Skills of Preschool Children with Autism Spectrum Disorders." *Journal of Autism and Developmental Disorders* 39(2): 231–41. http://www.ncbi.nlm.nih.gov/pubmed/18629623 (July 14, 2017).

Jones, Warren, and Ami Klin. 2013. "Attention to Eyes Is Present but in Decline in 2–6-Month-Old Infants Later Diagnosed with Autism." *Nature* 504(7480): 427–31. http://www.nature.com/doifinder/10.1038/nature12715 (November 27, 2017).

Jou, R. J. et al. 2009. "Brainstem Volumetric Alterations in Children with Autism." *Psychological medicine* 39(8): 1347–54. http://www.ncbi.nlm.nih.gov/pubmed/18812009 (December 15, 2016).

Jou, Roger J. et al. 2013. "A Two-Year Longitudinal Pilot MRI Study of the Brainstem in Autism." *Behavioural Brain Research* 251: 163–67. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4107420/pdf/nihms471503.pdf (May 11,

2017).

Katuwal, Gajendra J. et al. 2016. “Inter-Method Discrepancies in Brain Volume Estimation May Drive Inconsistent Findings in Autism.” *Frontiers in Neuroscience* 10. http://journal.frontiersin.org/article/10.3389/fnins.2016.00439.

Kleiman, M D, S Neff, and N P Rosman. 1992. “The Brain in Infantile Autism: Are Posterior Fossa Structures Abnormal?” *Neurology* 42(4): 753–60. http://www.ncbi.nlm.nih.gov/pubmed/1565227 (December 15, 2016).

Klein, Arno, and Jason Tourville. 2012. “101 Labeled Brain Images and a Consistent Human Cortical Labeling Protocol.” *Frontiers in Neuroscience* 6(DEC): 171. http://journal.frontiersin.org/article/10.3389/fnins.2012.00171/abstract (April 4, 2017).

Lai M-C, Lombardo MV, Suckling J, Ruigrok ANV, Chakrabarti B, Ecker C, Deoni SCL, Craig MC, Murphy DGM, Bullmore ET et al: Biological sex affects the neurobiology of autism. Brain 2013, 136(9):2799-2815.

Lalani SJ, Duffield TC, Trontel HG, Bigler ED, Abildskov TJ, Froehlich A, Prigge MBD, Travers BG, Anderson JS, Zielinski BA, Alexander A, Lange N, Lainhart JE. Auditory attention in autism spectrum disorder: An exploration of volumetric magnetic resonance imaging findings. J Clin Exp Neuropsychol. 2018 Jun;40(5):502-517.

Lambert, Christian et al. 2013. “Multiparametric Brainstem Segmentation Using a Modified Multivariate Mixture of Gaussians.” *NeuroImage: Clinical* 2(1): 684–94. http://www.sciencedirect.com/science/article/pii/S2213158213000569 (November 14, 2017).

Lord, Catherine et al. 2000. “The Autism Diagnostic Observation Schedule—Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism.” *Journal of Autism and Developmental Disorders* 30(3): 205–23. http://link.springer.com/10.1023/A:1005592401947 (October 27, 2016).

Lundwall, Rebecca A. et al. 2017. “Relationship between Brain Stem Volume and Aggression in Children Diagnosed with Autism Spectrum Disorder.” *Research in Autism Spectrum Disorders* 34: 44–51. http://www.sciencedirect.com/science/article/pii/S1750946716301556 (September 19, 2017).

Malone, Ian B. et al. 2015. “Accurate Automatic Estimation of Total Intracranial Volume: A Nuisance Variable with Less Nuisance.” *NeuroImage* 104: 366–72. http://www.sciencedirect.com/science/article/pii/S1053811914007769 (May 10, 2017).

Marcus, Daniel S. et al. 2007. “Open Access Series of Imaging Studies (OASIS): Cross-Sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults.” *Journal of cognitive neuroscience* 19(9): 1498–1507. http://www.ncbi.nlm.nih.gov/pubmed/17714011 (April 4, 2017).

Di Martino, Adriana et al. 2011. “Aberrant Striatal Functional Connectivity in Children with Autism.” *Biological psychiatry* 69(9): 847–56. http://www.ncbi.nlm.nih.gov/pubmed/21195388 (September 19, 2017).

Merker, Bjorn. 2007. “Consciousness without a Cerebral Cortex: A Challenge for Neuroscience and Medicine.” *Behavioral and Brain Sciences* 30(1): 63-81-134. http://www.ncbi.nlm.nih.gov/pubmed/17475053 (November 27, 2017).

Ozonoff, Sally et al. 2008. “The Onset of Autism: Patterns of Symptom Emergence in the First Years of Life.” *Autism research : official journal of the International Society for Autism Research* 1(6): 320–28. http://www.ncbi.nlm.nih.gov/pubmed/19360687 (July 14, 2017).

Panksepp, Jaak, and Lucy. Biven. 2012. *The Archaeology of Mind : Neuroevolutionary Origins of Human Emotions*. W.W Norton. https://books.google.it/books?id=bVdxXN_vVGEC&printsec=frontcover&dq=archaeology+of+mind:+Neuroevolutionary+origins+of+human+emotions&hl=it&sa=X&ved=0ahUKEwi_-

q71s97XAhVOF8AKHcFKBEgQ6AEIKjAA#v=onepage&q=archaeology of mind%3A Neuroevolutionary origins of (November 27, 2017).

Patenaude, Brian, Stephen M. Smith, David N. Kennedy, and Mark Jenkinson. 2011. "A Bayesian Model of Shape and Appearance for Subcortical Brain Segmentation." *NeuroImage* 56(3): 907–22. http://www.ncbi.nlm.nih.gov/pubmed/21352927 (April 4, 2017).

Pelphrey, Kevin A., James P. Morris, and Gregory McCarthy. 2005. "Neural Basis of Eye Gaze Processing Deficits in Autism." *Brain* 128(5): 1038–48. http://www.ncbi.nlm.nih.gov/pubmed/15758039 (November 27, 2017).

Porges, Stephen W. 2011. *The Polyvagal Theory: Neuropsysiological Foundations of Emotions, Attachment, Comunication and Self-Reguletation*. 1st ed. New York: W.W. Norton. https://books.google.it/books/about/The_Polyvagal_Theory_Neurophysiological.html?id=Wfa5DBL66aoC&redir_esc=y (November 27, 2017).

Porges, Stephen W., and Senta A. Furman. 2011. "The Early Development of the Autonomic Nervous System Provides a Neural Platform for Social Behaviour: A Polyvagal Perspective." *Infant and Child Development* 20(1): 106–18. http://www.ncbi.nlm.nih.gov/pubmed/21516219 (November 27, 2017).

Qiu, Ting et al. 2016. "Two Years Changes in the Development of Caudate Nucleus Are Involved in Restricted Repetitive Behaviors in 2-5-Year-Old Children with Autism Spectrum Disorder." *Developmental Cognitive Neuroscience* 19: 137–43. http://www.ncbi.nlm.nih.gov/pubmed/26999477 (September 19, 2017).

Retico, Alessandra et al. 2016. "The Effect of Gender on the Neuroanatomy of Children with Autism Spectrum Disorders: A Support Vector Machine Case-Control Study." *Molecular autism* 7(1): 5. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4717545&tool=pmcentrez&rendert

ype=abstract.

Rochat, Magali J et al. 2013. “Impaired Vitality Form Recognition in Autism.” *Neuropsychologia* 51(10): 1918–24. http://linkinghub.elsevier.com/retrieve/pii/S0028393213001802 (November 27, 2017).

Roid, GH, and LJ Miller. 1997. “Leiter International Performance Scale—revised (Standardized Edition).” *Wood Dale, IL USA: Stoetling Co*. https://scholar.google.com/scholar_lookup?title=The Leiter International Performance Scale-revised Edition&author=GH. Roid&author=LJ. Miller&publication_year=1997 (May 10, 2017).

Schumann CM, Barnes CC, Lord C, Courchesne E. Amygdala enlargement in toddlers with autism related to severity of social and communication impairments. Biol Psychiatry. 2009;66:942–9.

Sclocco, Roberta et al. 2017. “Challenges and Opportunities for Brainstem Neuroimaging with Ultrahigh Field MRI.” *NeuroImage*. http://www.ncbi.nlm.nih.gov/pubmed/28232189 (September 19, 2017).

Shen, Li et al. 2006. “Morphometric Analysis of Genetic Variation in Hippocampal Shape in Mild Cognitive Impairment: Role of an IL-6 Promoter Polymorphism.” *Technical Report*: 1–6. http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.118.7534 (April 21, 2017).

Shen, Li, Hany Farid, and Mark A McPeek. 2009. “Modeling Three-Dimensional Morphological Structures Using Spherical Harmonics.” *Evolution; international journal of organic evolution* 63(4): 1003–16. http://www.ncbi.nlm.nih.gov/pubmed/19154365 (April 21, 2017).

Styner, Martin et al. 2006. “Framework for the Statistical Shape Analysis of Brain Structures Using SPHARM-PDM.” *The insight journal* (1071): 242–50. http://www.ncbi.nlm.nih.gov/pubmed/21941375 (April 21, 2017).

Sutera, Saasha et al. 2007. “Predictors of Optimal Outcome in Toddlers Diagnosed with Autism

Spectrum Disorders." *Journal of autism and developmental disorders* 37(1): 98–107. http://www.ncbi.nlm.nih.gov/pubmed/17206522 (July 14, 2017).

Teitelbaum, P et al. 1998. "Movement Analysis in Infancy May Be Useful for Early Diagnosis of Autism." *Proceedings of the National Academy of Sciences of the United States of America* 95(23): 13982–87. http://www.ncbi.nlm.nih.gov/pubmed/9811912 (July 14, 2017).

Travers, Brittany G et al. 2015. "Brainstem White Matter Predicts Individual Differences in Manual Motor Difficulties and Symptom Severity in Autism." *Journal of Autism and Developmental Disorders* 45(9): 3030–40. http://www.ncbi.nlm.nih.gov/pubmed/26001365 (September 19, 2017).

Trevarthen, Colwyn, and Jonathan T Delafield-Butt. 2013. "Autism as a Developmental Disorder in Intentional Movement and Affective Engagement." *Frontiers in integrative neuroscience* 7: 49. http://www.ncbi.nlm.nih.gov/pubmed/23882192 (November 27, 2017).

Vandekerckhove, Marie, and Jaak Panksepp. 2011. "A Neurocognitive Theory of Higher Mental Emergence: From Anoetic Affective Experiences to Noetic Knowledge and Autonoetic Awareness." *Neuroscience and Biobehavioral Reviews* 35(9): 2017–25. http://www.sciencedirect.com/science/article/pii/S0149763411000674?via%3Dihub (November 27, 2017).

Venkatraman, Anand, Brian L Edlow, and Mary Helen Immordino-Yang. 2017. "The Brainstem in Emotion: A Review." *Frontiers in Neuroanatomy* 11: 15. http://www.ncbi.nlm.nih.gov/pubmed/28337130 (September 19, 2017).

Wang, Hongzhi et al. 2013. "Multi-Atlas Segmentation with Joint Label Fusion." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35(3): 611–23. http://www.ncbi.nlm.nih.gov/pubmed/22732662 (April 4, 2017).

Wechsler, D. 1974. "WISC-R, Wechsler Intelligence Scale for Children, Revised."

https://scholar.google.com/scholar_lookup?title=Wechsler Intelligence Scale for Children-Revised (WISC-R)&author=D. Wechsler&publication_year=1986 (May 10, 2017).

Wechsler, D, PI Scales, and VC Index. 2012. "Wechsler Preschool and Primary Scale of Intelligence—Fourth Edition." http://ux1.eiu.edu/~glcanivez/Adobe pdf/Publications-Papers/Canivez (2014) Buros MMY WPPSI-IV Review.pdf (May 10, 2017).

Welsh, John P., Edward S. Ahn, and Dimitris G. Placantonakis. 2005. "Is Autism due to Brain Desynchronization?" *International Journal of Developmental Neuroscience* 23(2–3 SPEC. ISS.): 253–63. http://www.sciencedirect.com/science/article/pii/S0736574804001236?via%3Dihub (November 27, 2017).

Worsley, KJ et al. 2009. "SurfStat: A Matlab Toolbox for the Statistical Analysis of Univariate and Multivariate Surface and Volumetric Data Using Linear Mixed Effects Models and Random Field Theory." *NeuroImage* 47(47): S102. http://linkinghub.elsevier.com/retrieve/pii/S1053811909708821 (April 21, 2017).

Figures:

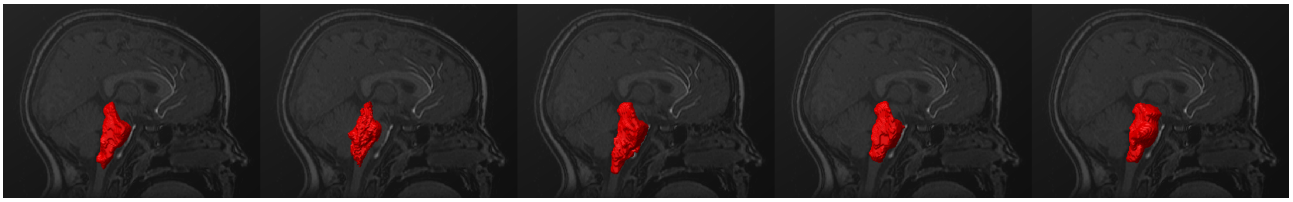

Figure 1: 3D renderings of brainstem segmentations of the same subject for the different methods (ANTs, FSL-FIRST, FreeSurfer 5.3, FreeSurfer 6.0, FreeSurfer 6.0 with substructures, respectively), overlaid on the corresponding median sagittal plane of the T1-MRI.

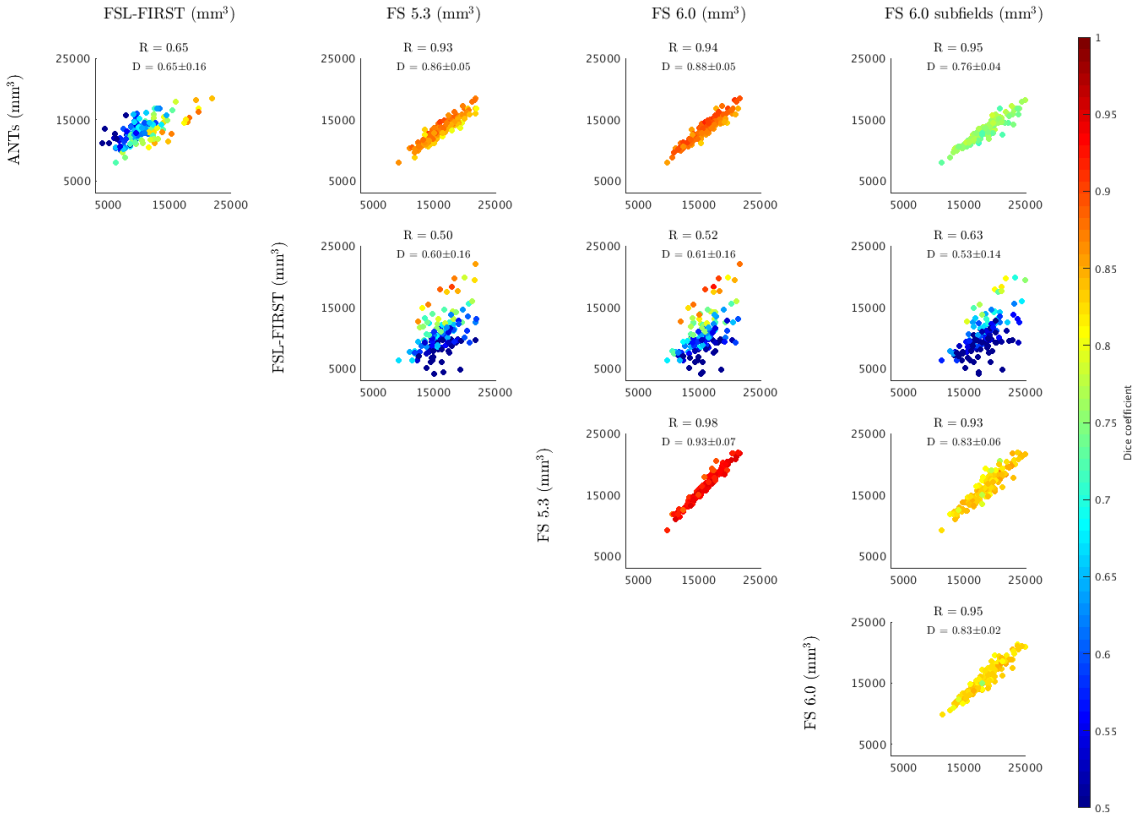

Figure 2: Scatter plots of the brainstem volumes extracted by different methods. The Pearson correlation coefficient (R) between the volumes obtained through each pair of segmentation methods is reported. The colours of the dots represent the Dice coefficient (D) computed for each pair of segmented masks in native space.

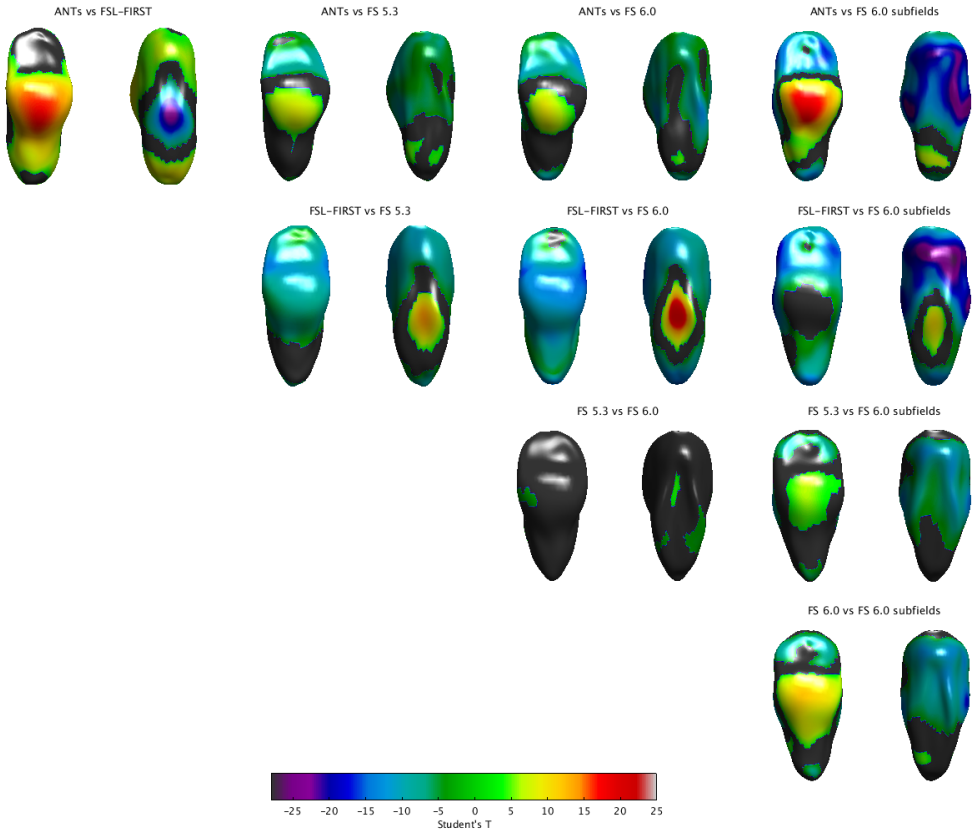

Figure 3: Shape differences between segmentation methods represented by T-maps of the distances along normal direction from a reference surface. The threshold is at cluster-level p-value=0.05 RFT corrected.

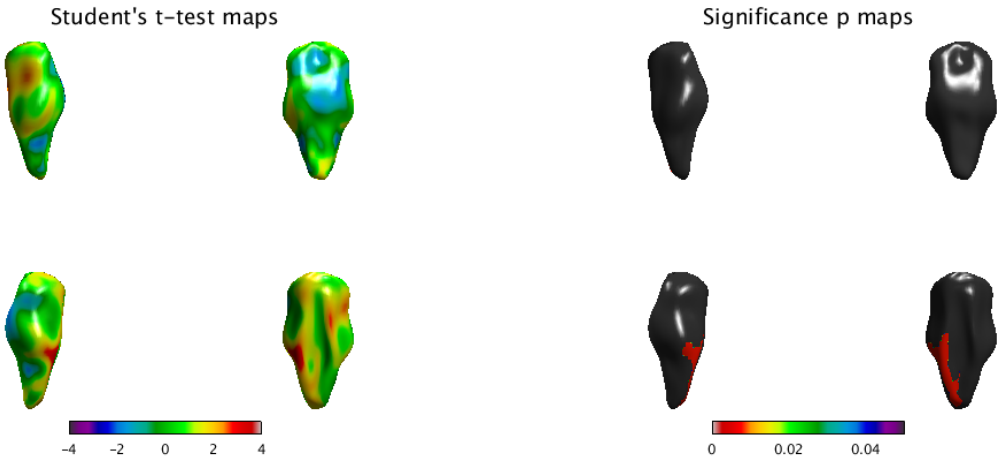

Figure 4a: Shape differences between ASD and Control group segmentations represented by T-maps of the distances along normal direction from the reference surface (for each map the left, anterior, posterior and right views are reported in clockwise order). The threshold is at cluster-level p-value=0.05 RFT corrected.

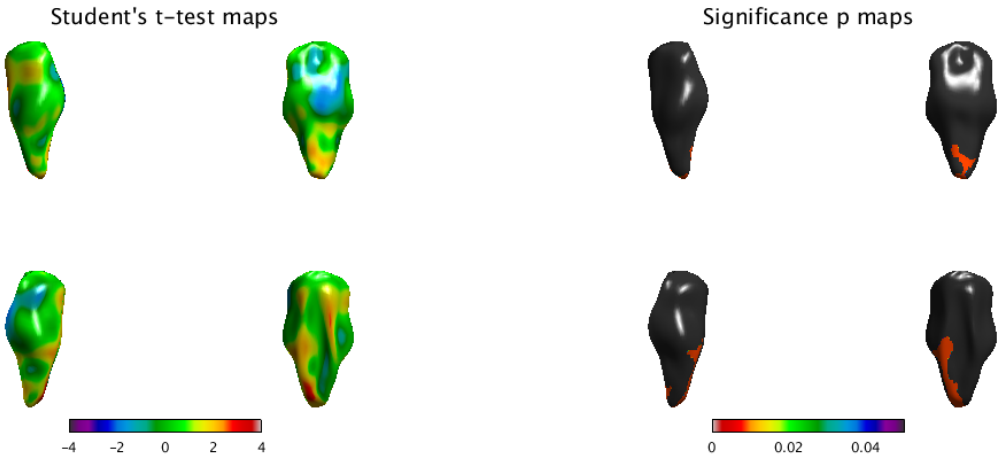

Figure 4b: Shape differences between ASD-ID and Control-ID group segmentations represented by T-maps of the distances along normal direction from the reference surface (for each map the left, anterior, posterior and right views are reported in clockwise order). The threshold is at cluster-level p-value=0.05 RFT corrected.

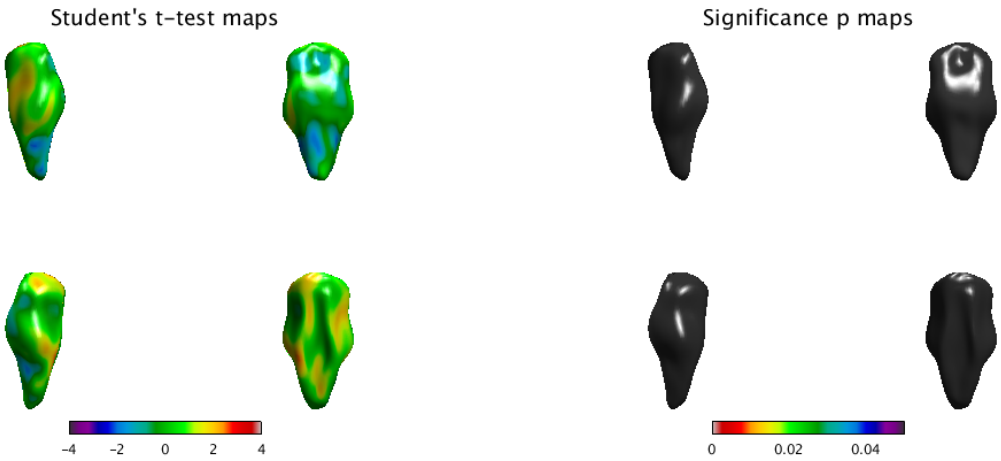

Figure 4c: Shape differences between ASD-no-ID and Control-no-ID group segmentations represented by T-maps of the distances along normal direction from the reference surface (for each map the left, anterior, posterior and right views are reported in clockwise order). The threshold is at cluster-level p-value=0.05 RFT corrected.

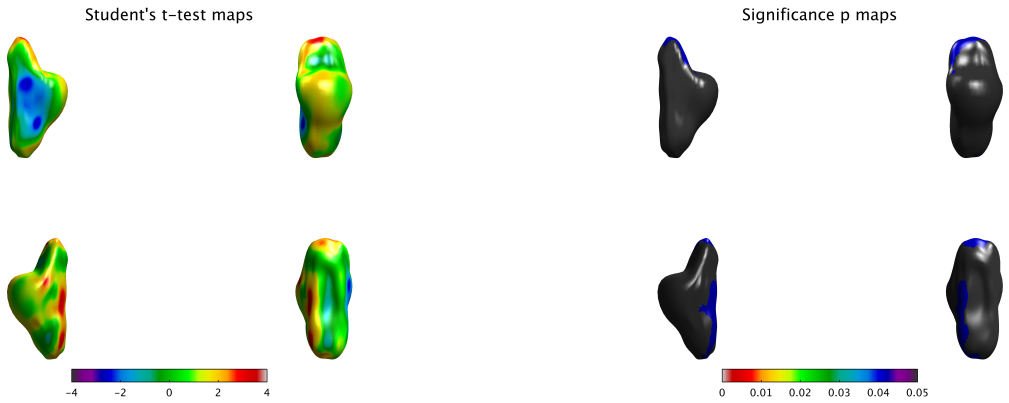

Figure 5a Annex A: Shape differences between ASD and Control group ANTS segmentations represented by T-maps of the distances along normal direction from the reference surface (for each map the left, anterior, posterior and right views are reported in clockwise order). The threshold is at cluster-level p-value=0.05 RFT corrected.

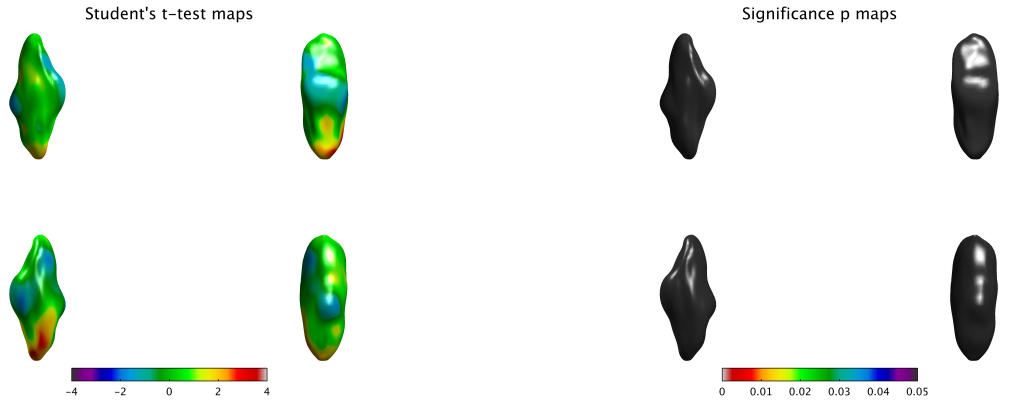

Figure 5b Annex A: Shape differences between ASD and Control group FSL-FIRST segmentations represented by T-maps of the distances along normal direction from the reference surface (for each map the left, anterior, posterior and right views are reported in clockwise order). The threshold is at cluster-level p-value=0.05 RFT corrected.

Annex A

Analogously to the FS with substructures, the shape comparison between ASD and typically developing children was also performed with ANTS and FSL-FIRST brainstem segmentations.

Figure 5a reports the results with ANTS segmentations. As in FreeSurfer case, the shape analysis confirms the enlargement of the brainstem segmented by ANTS, suggesting that the main differences are located in the posterior side, involving the medulla oblongata and the cerebellar peduncles structures.

Figure 5b reports the results with FSL-FIRST segmentations. Interestingly, the T-map shows a warm area in the same position in the medulla oblongata structure as with the other methods. However, the inclusion of the 4th ventricle in the segmentation protocol seems to hinder the detection of the differences and no significant areas with p-value < 0.05 RFT corrected could be identified.