High-resolution magnetic resonance imaging reveals nuclei of the human amygdala: manual segmentation to automatic atlas

Saygin Z. M.*[1,2] & Kliemann D.*[1,2], Iglesias J.E[3,4]. van der Kouwe, A.J.W.[2], Boyd E.[2], Reuter M[2]., Stevens A[2]., Van Leemput, K[2,5], McKee, A[6,] Frosch, M. P.[7], Fischl B.[2,8], Augustinack J. C [2]., for the Alzheimer's Disease Neuroimaging Initiative[+].

1. Massachusetts Institute of Technology/ McGovern Institute
43 Vassar St. Cambridge, MA 02139
2. Athinoula A Martinos Center, Dept. of Radiology, Massachusetts General Hospital, 149 13th Street, Charlestown MA 02129 USA
3. University College London, Dept. Medical Physics and Biomedical Engineering Translational Imaging Group, Malet Place Engineering Building, Gower Street, London, WC1E 6BT, UK
4. Basque Center on Cognition, Brain and Language, Paseo Mikeletegi 69, 20009 Donostia - San Sebastian, Spain
5. Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark
6. Department of Neurology and Pathology, Boston University School of Medicine, Boston University Alzheimer's Disease Center, Boston MA 02118, VA Boston Healthcare System, MA 02130 USA
7. C.S. Kubik Laboratory for Neuropathology, Pathology Service, MGH, 55 Fruit St., Boston MA 02115 USA
8. MIT Computer Science and AI Lab, Cambridge MA 02139 USA

* Authors contributed equally to this work.

[+]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

ABSTRACT

The amygdala is composed of multiple nuclei with unique functions and connections in the limbic system and to the rest of the brain. However, standard *in vivo* neuroimaging tools to automatically delineate the amygdala into its multiple nuclei are still rare. By scanning postmortem specimens at high resolution (100-150μm) at 7T field strength (n = 10), we were able to visualize and label nine amygdala nuclei (anterior amygdaloid, cortico-amygdaloid transition area; basal, lateral, accessory basal, central, cortical medial, paralaminar nuclei). We created an atlas from these labels using a recently developed atlas building algorithm based on Bayesian inference. This atlas, which will be released as part of FreeSurfer, can be used to automatically segment nine amygdala nuclei from a standard resolution structural MR image. We applied this atlas to two publicly available datasets (ADNI and ABIDE) with standard resolution T1 data, used individual volumetric data of the amygdala nuclei as the measure and found that our atlas i) discriminates between Alzheimer's disease participants and age-matched control participants with 84% accuracy (AUC=0.915), and ii) discriminates between individuals with autism and age-, sex- and IQ-matched neurotypically developed control participants with 59.5% accuracy (AUC=0.59). For both datasets, the new *ex vivo* atlas significantly outperformed (all $p < .05$) estimations of the whole amygdala derived from the segmentation in FreeSurfer 5.1 (ADNI: 75%, ABIDE: 54% accuracy), as well as classification based on whole amygdala volume (using the sum of all amygdala nuclei volumes; ADNI: 81%, ABIDE: 55% accuracy). This new atlas and the segmentation tools that utilize it will provide neuroimaging researchers with the ability to explore the function and connectivity of the human amygdala nuclei with unprecedented detail in healthy adults as well as those with neurodevelopmental and neurodegenerative disorders.

Highlights:

-We visualized 9 nuclei boundaries (anterior amygdaloid area, cortico-amygdaloid transition area; basal, lateral, accessory basal, central, cortical medial, paralaminar nuclei) using ultra-high-resolution *ex vivo* imaging

-Nuclei were consistent across cases and raters

- We built a segmentation atlas of the amygdala nuclei, which will be distributed with FreeSurfer
-The atlas was applied to 2 separate datasets and demonstrated higher discriminability of Alzheimer's disease & autism than previously possible with amygdala segmentation methods
-The atlas will provide neuroimaging researchers with the ability to test nucleus function with greater spatial specificity

Keywords (6 max): amygdala, medial temporal lobe, atlas, *ex vivo*, autism, Alzheimer's

1. INTRODUCTION

The amygdala is composed of heterogeneous nuclei, defined primarily by their distinct cytoarchitecture, neurotransmitters, and connectivity patterns (Freese and Amaral, 2005, 2006, 2009; Alheid, 2003; Price et al., 1987; Aggleton, 2000; Gloor, 1972, 1978, 1997; McDonald, 1998; LeDoux, 1998, De Olmos, 2004; De Olmos & Heimer, 1999). Studies on rodents and non-human primates have advanced our understanding of the functions of the individual nuclei. For example, the lateral (La) and basal (Ba) nuclei are engaged in updating current stimulus value associations, primarily through connections with orbitofrontal regions (Baxter and Murray, 2002); the central nucleus (Ce) is believed to mediate behavioral responses to potentially harmful stimuli and fear perception through its connectivity with hypothalamus, basal forebrain, and the brainstem (Kalin et al., 2004; Phillips & LeDoux, 1992). In humans, the amygdala as a whole is thought to play a key role in emotional and social cognitive processes (e.g. Adolphs et al., 2005, Kliemann et al. 2012, Hortensius et al., 2016), and accordingly, its dysfunction is implicated in psychopathologies, such as mood disorders (Phillips et al., 2003; Siegle et al., 2002), anxiety disorders (Birbaumer et al., 1998; Rauch et al., 2003), and developmental disorders (Baron-Cohen et al., 2000; Dziobek et al., 2010). Additionally, several post mortem studies have shown that the amygdala is a common site for neurofibrillary tangles and senile plaques in mild cognitive impairment (Merkesbery, 2010) and Alzheimer's disease (Yilmazer-Hanke, 1998) as well as Lewy bodies (Kotzbauer et al., 2011, Fujishiro et al., 2002).

However, the relationship between the structure and function of the distinct nuclei in humans remain largely unknown, both in health and disease. The small size of the amygdala's nuclei has made it difficult to study this structure noninvasively in the living brain using standard neuroimaging resolution. Previous segmentation studies of the amygdala have used either i) visual approximation based on a single-subject histological atlas (Etkin et al., 2004; dorsal vs. ventral amygdala Dolan, 2002, 2007); ii) manual segmentations based on *in vivo* neuroimaging; iii) normalization and application of a probabilistic atlas (Amunts et al., 2005, Solano-Castiella et al., 2011); or iv) segmentations based on diffusion-weighted imaging. The first two approaches are labor intensive and susceptible to human error. Using the reference space of the MNI single subject, has limited applicability in segmentation for two reasons: first, spatial normalization can lead to inaccuracies due to the fact that the annotations were made on histology, which leads to blurry probability maps; and second, the direct warping of such probability maps to obtain segmentations greatly suffers from registration errors. Additionally, these previous approaches have segmented the amygdala into 2-4 nucleus groups. The use of diffusion- weighted imaging to segment the amygdala has been attractive due to the possibility of automation and within-subject segmentation (rather than normalization to a template). Fiber orientations within the amygdala have been used to divide the structure into two subregions, centromedial and basolateral (Solano-Castiella et al., 2010). However, this method, like others before it, performed analyses on images normalized to a template brain, and were restricted to only two subdivisions. Diffusion connectivity patterns have also been used to delineate each individual's amygdala into four nucleic groups, using nucleus-specific connectivity patterns based on previous animal literature (Saygin et al. 2011; Saygin et al. 2015). While this method offered more nucleic groups (parcellated into 4 groups), the nuclei were dependent on each individual's connectivity patterns, which may be compromised in some patient populations. Thus, a segmentation method independent of connectivity and with finer detail (i.e. nuclei instead of subregions) offers a better understanding of the individual nuclei of the amygdala.

Without an easily accessible technique with which to parcellate the amygdala, it is difficult to elucidate the separate roles of the human amygdala nuclei, as well as the impact of individual differences in nucleus structure and function. Moreover, progress towards mechanistic theories of dysfunction and abnormal development will remain hindered until these structures can be explored *in vivo*.

Here, we use *ex vivo* MRI data from autopsy brains to delineate the amygdala nuclei and build a probabilistic atlas of amygdala anatomy, using a novel algorithm, which will be distributed as part of the FreeSurfer software. Our *ex vivo* imaging protocol yields images with extremely high resolution and signal-to-noise ratio, dramatically higher than is possible *in vivo*, which allows us to accurately identify more nuclei with a segmentation protocol specifically designed for this study. We were able to define nine amygdala nuclei that are major subdivisions in human and animal histology literature (e.g. deOlmos 2004; Gloor et al., 1997; Brockhaus 1938; Sims & Williams, 1990; Freese & Amaral, 2009; Whalen & Phelps 2009; LeDoux 1998), and whose boundaries are clearly visible in the *ex vivo* images (see also **Methods**). This segmentation focuses on the main amygdala nuclei in the medial temporal lobe and not the extended amygdala. Our previous work - the *ex vivo* hippocampal atlas (Iglesias et al., 2015) - uses a generative modeling framework to directly segment individual subject *in vivo* MRI data in target space; the resulting segmentation algorithm can be used to analyze standard *in vivo* MRI scans with varying overall image contrast properties and intensity distributions, while producing sharper and more accurate label posterior probabilities than direct registration to a reference space. Here, we use this approach and extend it to the amygdala. We also apply this atlas to two publicly available datasets with standard resolution T1 data, and evaluate how well the resulting amygdala nucleus segmentation volumes can classify i) individuals with Alzheimer's disease and older adult controls and ii) individuals with autism and age-matched controls.

2. MATERIALS and METHODS

2.1 Autopsy brain samples and *ex vivo* MRI acquisition

The dataset of *ex vivo* scans comprised 10 autopsied brain hemispheres from the Massachusetts General Hospital Autopsy Service (Massachusetts General Hospital, Boston, MA) and from the Framingham Heart Study and Boston University Alzheimer's Disease Center (Veterans Administration Medical Center, Bedford, MA). Samples consisted of 5 right and 5 left hemispheres (or blocks encompassing the amygdala) of 10 cases (9 without any neurological conditions, 1 with mild AD). Table 1 lists the subject-specific demographic information. In short, subjects were on average 67 years old at the time of death, 2 were female, and the post-mortem interval did not exceed 24 hours. Please note that we use the term 'case' to refer to hemispheres. We use this terminology to ensure that each case represents one hemisphere from a separate individual (not, e.g. 10 hemispheres from 5 individuals).

From each *ex vivo* sample, a block of tissue surrounding the amygdala (or the complete MTL) was excised and first fixed with 10% formalin and then transferred to periodate–lysine–paraformaldehyde (PLP). Depending on its size, the block was placed in either a plastic cylindrical centrifuge tube (60 ml, 3 cm diameter) or inside a plastic sealed bag filled with PLP. In the latter case, air was pumped out using a needle and a vacuum pump in order to minimize the number and size of air bubbles in the samples. The tissue block was subsequently scanned in a 7 T Siemens scanner using a 3D FLASH sequence with TR = 60msec, TE = 30 msec, $\alpha = 20°$ (Fischl et al., 2009; Augustinack et al., 2013). Six of the samples were scanned at 0.1 mm isotropic resolution (three with TE = 12.8 msec and TR = 40msec; and three with TE1 = 10.75 msec, TE2 = 25.5 msec and TR = 45 msec $\alpha = 35°$), three at 0.12 mm and one at 0.15 mm. Radio frequency coils were used in the acquisition, accommodating variations in sample size: either a 4-turn solenoid coil (28.5 mm inner diameter, 44 mm length), a 4-channel phased-array (a linear array of loop coil elements each with 5 cm coil diameter, 1.5 cm overlap between adjacent elements, 16 cm in length) were used. Despite the fact that different coils were used to scan the different samples, the output images were comparable in quality. The whole procedure received IRB approval before its execution by the Partners Human Research Committee; thus, all tissue was collected in accordance with approved protocols.

Manual segmentation of *ex vivo* MRI data: anatomical definitions

J.C.A., Z.M.S., and D.K. developed the amygdala segmentation protocol based on several histology and morphometry resources (deOlmos 2004; Gloor et al., 1997; Brockhaus 1938; Sims & Williams, 1990; we also surveyed the monkey amygdala parcellation (Freese & Amaral, 2009), Whalen & Phelps 2009, Chapter 1) as an additional guide but focused on the human literature as the primary source for our parcellation. The amygdala annotations are described in Table 2: lateral nucleus, basal nucleus, accessory basal nucleus, central nucleus, medial nucleus, cortical nucleus, anterior amygdaloid area, cortico-amygdaloid transition area. In humans, CAT is the equivalent to the periamygdaloid cortex (PAC) in animals, except layer PAC-3 which this protocol considers as part of the cortical nucleus (see below). Although several other nuclei of the amygdala have been described in histologic preparations (e.g. intercalated nuclei or subdivisions of each nuclei), we only labeled nuclei that were visible via *ex vivo* MRI contrast. Note that the descriptions are not based on histology but mainly on the contrast at the boundary that was visible in the *ex vivo* MRI data. We describe the manual labeling protocol in detail based on the contrast visible in the *ex vivo* data.

Three manual raters (Z.M.S., D.K., and E.B.) then applied the developed protocol to the ten *ex vivo* cases independently. Each case required several weeks (average 4 weeks) to annotate and varied slightly with image quality (due to human brain tissue variability) and in resulting image contrast. Although the overall image contrast varied across the cases, each case had sufficient image quality to be able to determine the boundary between nuclei and label these nuclei based on this contrast at the boundaries. Due to the high resolution of the scans, the resulting number of slices containing amygdala in coronal orientation was 133 (SD: 28.1) while a standard *in vivo* T1 (1mm resolution) typically contains ~18 slices of amygdala. We used Freeview, a visualization tool implemented in FreeSurfer to perform the manual labeling of the amygdala nuclei. Most MR volumes needed a slight rotation, which align the brain tissue to a coronal view and standardized the orthogonal planes for labeling. In the rare case of bubbles or tissue damage on MR images (2 cases, spreading only a very small portion of all slices), we labeled affected voxels according to i) the nuclei surrounding unaffected tissue as well as ii) the appearance of the nuclei anterior to posteriorly.

In general, annotations were made in the coronal view. Axial and sagittal views were used in addition to guide delineation of subregions and these additional views (and the 3D nature of MRI) helped better define the nuclei contrast and borders. Note that the strategy to annotate mainly in coronal view leads to slightly jagged boundaries in the two other views. However, this effect is averaged out during downsampling and construction of the atlas. To maximize the high level of anatomical consistency across raters, J.C.A. served as quality control, while supervising and suggesting refinement of delineations if necessary based on MRI boundary contrast between nuclei. The objective of the neuroanatomist was to verify that the labeling was true to the MRI contrast to ensure quality labeling and not to influence the labeling itself.

Atlas construction

In this study, we encode the anatomical variability of the amygdala and surrounding tissue into a statistical atlas. Following Van Leemput (2009), the atlas is represented by a tetrahedral mesh that covers the amygdala (and surrounding structures) in a canonical space. Each vertex in the mesh has an associated vector of label probabilities, which contains the relative frequencies with which each neuroanatomical label is observed at each location. These probabilities are estimated at non-vertex locations using barycentric interpolation. The mesh is endowed with a deformation model, which allows it to cover the spectrum of anatomical variability in the population of the training data. This model infinitely penalizes the folding or collapsing of tetrahedra, which effectively preserves the topology of the mesh when deformed (Ashburner et al., 2000).

Constructing the atlas requires estimating the label probabilities at each vertex and the topology of the mesh, given a number of training manual segmentations. Here we use a modified version of Van Leemput's algorithm (Iglesias et al., 2015) that enables us to integrate training datasets that are labeled with different protocols, such that we can combine the manual segmentations of the *ex vivo* data (nuclei of the amygdala) and the *in vivo* scans (whole amygdala and surrounding structures). These two datasets are complementary: the former informs the atlas on the internal structure of the amygdala,

whereas the latter provides information on the outer contour of the amygdala and its surrounding structures.

The training process is based on Bayesian inference, i.e., answering the question: "what was the statistical atlas that most likely generated the manual segmentations?" The algorithm amounts to a group-wise registration of the manual segmentations using a high-density tetrahedral mesh followed by a mesh simplification process. The simplification uses a Bayesian model selection algorithm that automatically encodes the label uncertainty (blurring) in each region: convoluted areas, which are well represented in the training data are covered by small tetrahedra, while flat areas are covered by larger tetrahedra. Further details of the atlas representation can be found in Van Leemput (2009) and Iglesias (2015). A coronal section of the atlas with and without the tetrahedral mesh is shown in **Figure 3**.

Segmentation of *in vivo* MRI data

The segmentation of the data is posed as a Bayesian inference problem within a generative model of MRI images. The model assumes that: i) the atlas is first warped according to its deformation model; ii) a segmentation is then sampled from the label probabilities; and iii) image intensities are sampled at each voxel from a Gaussian distribution whose mean and variance depend on the label (tissue type) of the voxel. Within this model, the estimation of the segmentation is again posed as a Bayesian inference problem: "Given the atlas and the image intensities, what is the most probable segmentation?"

Finding the most like segmentation requires the minimization of a cost function that consists of two terms: a prior that encodes the cost of deforming the atlas and a likelihood term related to the probability of observing the image intensities given the segmentation. The segmentation algorithm minimizes the cost function by alternately optimizing the deformation of the atlas mesh and the Gaussian parameters (means and variances). The fact that these means and variances are estimated directly from the MRI scan to analyze (rather than encoded in the prior) makes the algorithm robust to changes in MRI contrast. Further details can be found in Van Leemput (2009), Van Leemput (2009b) and Iglesias (2015).

In vivo training MRI data

In order to be useful in segmentation, our probabilistic atlas needs to describe not only the amygdala nuclei but also the statistical distribution of the surrounding anatomy (e.g., hippocampus, cortex, etc). This distribution is learned from a separate dataset of 39 *in vivo* brain T1 MRI scans (19 males, 20 females, mean age: 56.3 years, 29 controls, 10 mildly demented) that where manually labeled at the structure level (i.e., whole amygdala, whole hippocampus, etc). We note that this was the dataset that was used to estimate the probabilistic atlas in the main FreeSurfer "recon-all" stream (Fischl et al., 2002, Fischl et al., 2004). Using the technique described in Iglesias et al. (2015), we combined the *in vivo* and *ex vivo* delineations into a single probabilistic atlas, including both the amygdala nuclei and surrounding structures.

In vivo MR data (quantitative test sets)

We further tested whether the probabilistic atlas can not only consistently identify amygdala nuclei in individual standard anatomical *in vivo* scans, but also whether the information about nuclei can be used to reliably distinguish between neuropathologic and neurotypical groups. To this end, we compared two neuropathologic groups with matched control groups in which the expected between-group structural differences ranged from gross to subtle.

The first comparison was based on a dataset of MRI scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). AD shows substantial neuropathology in medial temporal lobe structures, including the amygdala.

As a second group comparison we chose a population that has a more diverse neuropathology in the amygdala – no clear consensus has emerged in autism yet. We compared age-, sex-, IQ-matched typically developed healthy control sample and Autism Spectrum Disorders (ASD). Data was taken from the Autism Brain Imaging Data Exchange initiative (ABIDE, http://fcon_1000.projects.nitrc.org/indi/abide/, DiMartino et al., 2014). Both datasets were processed with FreeSurfer version 5.1.

Information about the in vivo samples ADNI

The selection from the ADNI-dataset comprised 213 AD individuals and 161 healthy control participants matched for age (AD: 76.04 (SD 5.42), CNT: 75.58 (SD 7.37), $t(372) = .7$, $p = .48$). The ADNI was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public-private partnership. The main goal of ADNI is to test whether MRI, positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to analyze the progression of MCI and early AD. Markers of early AD progression can aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as decrease the time and cost of clinical trials. The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California — San Francisco. ADNI is a joint effort by co-investigators from industry and academia. Subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 subjects but ADNI has been followed by ADNI-GO and ADNI-2. These three protocols have recruited over 1500 adults (ages 55–90) to participate in the study, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow up duration of each group is specified in the corresponding protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see http://www.adni-info.org. MR images are T1-weighted and have 1mm isotropic resolution. Exact acquisition parameters depend on the site that acquired the data. Further details can be found in adni-info.org.

ABIDE

The selection from the ABIDE-dataset included 131 individuals on the Autism Spectrum and 131 neurotypically-developed control subjects and was a priori to analysis. The selection of data from the complete dataset was motivated as follows: First, a qualitative assessment of FreeSurfer derived amygdala segmentations was performed by visually inspecting 3D representations thereof. The resulting number of cases was further reduced by choosing only adults (age > 18) with an IQ over 90. Third, we consecutively

reduced the number of control subjects until the number of subjects in each group was equal by excluding control subjects with the highest IQ within their group. This procedure resulted in groups matched for age (ASD: 26.4 (SD 8.88), NT: 25.95 (SD: 6.99), $t(26) = .43$, $p = .67$), FIQ (ASD: 111.87, NT: 11179, $t(260) = .07$, $p = .95$) and sex (ASD: 16 females, NT: 18 females).

Discrimination analyses

In this work, we use the ability to discriminate groups using volumes of nuclei as a proxy for segmentation quality. In order to ensure that the accuracy of the discrimination is mostly determined by the quality of the input features (i.e., volumes of nuclei) rather than fluctuations in the classifier, we use a simple linear discriminant analysis (LDA, Fisher 1936). More specifically, we use a leave one out scheme in which, for each subject, we first compute the direction that best separates the two classes using all other subjects, and then evaluate the projection of the subject at hand on that direction to compute a scalar score. Once we have the scores for all the subjects, we carry out two analyses. First, we compute the p-value of a non-parametric statistical test (Wilcoxon rank sum) comparing the scores of the two groups. And second, we build a receiver operating characteristic (ROC) curve, and compute the discrimination accuracy at the optimal point of operation ("elbow"), as well as the area under the curve (AUC). In order to statistically compare the performance of two atlases, we use a paired DeLong test (DeLong et al., 1998) that compares the areas under the corresponding ROC curves. This comparison tests whether information derived from the new atlas, either using the sum of all nuclei or using all the nuclei volumes as a multi-dimensional input i.e., simultaneously with LDA will outperform an existing segmentation of the whole amygdala (FreeSurfer segmentation).

Analyses of hemisphere, sex and age on nucleus volume

We conducted further analyses to test for the influence of age, hemisphere and sex on nuclei volumes in the *in vivo* test sets (ABIDE and ADNI). We tested for hemispheric and sex differences with separate repeated measures ANOVAs with the between-subject

factor as *group* (ASD versus CNT) and the within-subject factors as *nucleus* (La, Ba, AB, Ce, Me, Co, CAT, AAA, PL), *hemisphere* (left, right), or *sex* (male, female). ICV was added as a covariate to all ANOVAs. To test for the influence of age on nucleus volume we conducted partial correlations with ICV as a covariate for both test datasets. We used Fisher's r to z transformations to test for differences in between-group correlations. Given the absence of hemispheric differences (see **Results**) nuclei volumes were averaged over both hemispheres. To control for multiple comparisons, p values were corrected for the number of nuclei i.e. an adjusted p value of .0055 (.050/9) in post-hoc tests or correlations.

RESULTS

Inter-rate reliability and volume of nuclei from *ex vivo* manual labeling

We labeled 10 postmortem cases based on the boundaries of nine nuclei that were clearly visible on the high resolution *ex vivo* images collected at 7 T (**Figure 1**). Inter-rater reliability was calculated as the Dice coefficient (overlapping voxels between the two independently labeled cases divided by the union of voxels) (Dice, 1945). Two individuals (Z.M.S., D.K., and E.B.) labeled case #1 in its entirety and the reliability was quite high: La 0.85; AB 0.76; Ba 0.73; Me 0.68; Ce 0.60; CAT 0.59; AAA 0.46; Co 0.44; PL 0.41; (**Figure 2**). All other cases were labeled in their entirety by one individual, and 18 slices (6 contiguous posterior, 6 contiguous middle, and 6 contiguous anterior slices) of each case were labeled by the other two labelers to calculate inter-labeler overlap measurements of all three labelers for each case (Dice coefficient mean ± se across all cases: La 0.83 ±0.02; AB 0.73 ±0.02; Ba 0.73 ±0.02; Me 0.34 ±0.06; Ce 0.54 ±0.03; CAT 0.61 ±0.03; AAA 0.38 ±0.06; Co 0.45 ±0.05; PL 0.35 ±0.04).

Coronal slices (**Figure 1**) were mainly used to determine labeling but sagittal and axial views were also used to visualize the borders and were especially useful for the Me, Ce, and Co nuclei due to their elongated shape in these views (**Figure 2**). The 3-dimensional rendering of one of the *ex vivo* cases illustrates the oblong versus spherical shapes of nuclei in different orientations, and also illustrates how intertwined some of these nuclei are (AB and Ce; CAT and Me; **Figure 4**). It also captures the differences in volume between the nuclei, with La, Ba, AB, and CAT occupying most of the amygdala's volume. **Table 3** details the mean volume of each nucleus for all fully-labeled cases (ten cases).

Atlas generation and applications to *in vivo* MRI

An atlas was generated from the post-mortem manual labels (see **Methods**) and applied to standard *in vivo* MR images in two public datasets: ADNI and ABIDE. These scans were acquired with MPRAGE sequences at 1 mm isotropic resolution. The MRI data were processed through the standard FreeSurfer pipeline (Fischl et al., 2002, Fischl et al., 2004), including the current automated amygdala segmentation, which is useful to

compare the segmentations yielded by the *in vivo* atlas of the whole amygdala with those produced by the *ex vivo* atlas we are introducing here. The resulting segmentations from the *ex vivo* atlas were not manually edited.

ADNI dataset:

We calculated how well the *ex vivo* based segmentations discriminate AD from control cases based on nucleus volume. Accuracy and area under the curve (AUC) were computed in a leave one out manner (all analyses were corrected for age, sex, and ICV). The *ex vivo* atlas was highly accurate at discriminating between AD and control participants; using the sum of the amygdala nuclei volumes as the discriminating feature yielded 81.46% accuracy (AUC=0.83; $p = 7.65 \times 10^{-41}$ AD vs. control), and significantly outperformed the whole-amygdala atlas in FreeSurfer v5.1 (Paired DeLong test for AUC of new atlas vs. FreeSurfer v5.1: $p = 1.8 \times 10^{-6}$). Using all the amygdala nuclei volumes simultaneously from the current study was also highly accurate in discriminating AD vs. controls, with 84.07% accuracy (AUC = 0.9154; $p = 2.80 \times 10^{-44}$) and offered significant improvement in discrimination as compared to i) using the previous FreeSurfer atlas (DeLong test $p = 9.\times10^{-6}$) and ii) using the sum of all nucleus volumes from the current atlas (DeLong test $p = 1.6 \times10^{-2}$) (see Table 4).

ABIDE dataset:

We also applied the *ex vivo* segmentation atlas to another public dataset of autism and control participants (ABIDE; **Figure 5**) and calculated discriminability based on nucleus volume. Neuroimaging differences and effect sizes are notoriously quite small in the ASD literature. We wanted to know i) how well amygdala nucleus volumes would be able to discriminate between ASD and control participants and ii) whether the discrimination performance would be significantly better than using the whole amygdala (the sum of all nuclei volumes derived by the atlas) and better than the whole amygdala volume from the standard FreeSurfer segmentation. Accuracy and AUC were computed in exactly the same way as the ADNI dataset; and note that we did not control for age as a nuisance regressor, given that groups were well matched on this variable. The previous FreeSurfer atlas failed to discriminate between ASD vs. controls ($p = 0.16$ with 54.05%

accuracy and AUC = 0.449; **Table 5**). In contrast, the *ex vivo* atlas was significantly accurate at discriminating between ASD and controls when using all the nuclei simultaneously (p = .0122), but not when using the sum of all the nuclei volumes (p = 0.078; see **Table 5** for accuracy and AUC). Using all nuclei simultaneously yielded 59.46% accuracy and AUC = 0.5902, offering substantial improvement in discrimination as compared to i) using the whole amygdala volume from FreeSurfer's automatic segmentation (DeLong test p = 1.7e-02) and ii) using the sum of all nucleus volumes from the current atlas (DeLong test p = 9.1e-03). The difference in AUC between the whole amygdala using FreeSurfer 5.1 or the aggregate volume of the nuclei was not significant (DeLong test p = 0.65). Therefore, despite the lack of contrast in the internal boundaries of the amygdala, the volumes of the nuclei carry additional information that is not present in the volume of the whole amygdala.

Additional analyses of hemisphere, sex, and age:

Hemispheric differences in nuclei volumes

The repeated measures ANOVAs (with a between-subject factor of *group*: ASD or ALZ vs. CNT, and within-subject factors of *hemisphere*: left vs. right and *nucleus*: La, Ba, AB, Ce, Me, Co, CAT, AAA, PL) showed no main effect of hemisphere or interactions of hemisphere with nucleus for both *in vivo* test datasets (main effect of hemisphere: ABIDE: p > .3, ADNI: p > .06; interactions with hemisphere and nucleus: ABIDE: p > .5, ADNI p > .23).

Although not central to our main question of whether nuclei are influenced by hemispheric differences, here we also report the full ANOVA results for the sake of completeness. For the ADNI dataset, there were main effects of nucleus ($F_{(1.3,380)} = 23.42$, $p = 1.7x10^{-7}$, $\eta = .058$, Greenhouse-Geisser corrected), group ($F_{(1.,380)} = 282.2$, $p = 9.2x10^{-48}$, $\eta = .43$) as well as a significant interaction of nucleus and group ($F_{(1.3,380)} = 193.2$, $p = 5.2x10^{-44}$, $\eta = .34$, Greenhouse-Geisser corrected), further replicating our initial prediction analyses above. The influence of the covariate was significant ($F_{(1.3,380)} = 98.6$, $p = 8.2x10^{-22}$, $\eta = .21$), and interacted with hemisphere ($F_{(1,380)} = 6.69$, $p = .01$, $\eta = .017$, Greenhouse-Geisser corrected) and nucleus ($F_{(1.3,380)} = 89.05$, $p = 3.1^{-23}$, $\eta = .19$, Greenhouse-Geisser corrected).

For the ABIDE dataset, there were main effects of nucleus ($F_{(1.4,256)} = 211.3$, $p = 6.5x10^{-47}$, $\eta = .45$, Greenhouse-Geisser corrected), the covariate ICV ($F_{(1,256)} = 101.96$, $p = 2.1x10^{-20}$, $\eta = .29$) and a significant interaction of nucleus and ICV ($F_{(1,256)} = 85.7$, $p = 1.5x10^{-122}$, $\eta = 85.6$, Greenhouse-Geisser corrected).

Sex differences in nuclei volumes

In the ADNI dataset, a repeated measures ANOVA (with between-subject factors of *group*: ASD vs. CNT and sex, within-subject factor *nucleus*, covariate ICV) revealed a significant main effect of sex ($F_{(1.3)} = 9.3$, $p = .003$, $\eta = .02$, Greenhouse-Geisser corrected), as well as a significant interaction of sex and nucleus ($F_{(1.3)} = 9.4$, $p = 9.8^{-4}$, $\eta = .02$, Greenhouse-Geisser corrected). Post-hoc one-way ANOVAs testing sex differences for each nucleus separately, while controlling for ICV as a covariate, showed differences between hemispheres only in La ($F_{(1.382)} = 11.6$, $p = .0018$, $\eta = .025$, corrected for multiple comparisons). All remaining nuclei showed trends towards greater nuclei volume in males (AAA: $p = .028$, AB: $p = .027$, Ba: $p = .020$, CAT: $p = .036$, Ce: .13, Me: $p = .055$, Co: $p = .017$, PL: $p = .022$).

There were main effects of nucleus ($F_{(1.3,380)} = 32.8$, $p = 5.8x10^{-10}$, $\eta = .08$, Greenhouse-Geisser corrected), group ($F_{(1.,380)} = 282.5$, $p = 9.2x10^{-48}$, $\eta = .43$) and a significant interaction of nucleus and group ($F_{(1.3,.380)} = 193.4$, $p = 3.1x10^{-44}$, $\eta = .34$, Greenhouse-Geisser corrected), again replicating our initial prediction analyses above. The influence of the covariate was significant ($F_{(1.3,380)} = 38.8$, $p = 1.3x10^{-9}$, $\eta = .09$), and interacted with nucleus ($F_{(1.3,380)} = 33.5$ $p = 3.8x10^{-10}$ $\eta = .081$, Greenhouse-Geisser corrected).

Given the very small number of females in the matched ABIDE dataset (ASD: n = 16, CNT: n = 18) we did not test the influence of sex in the ABIDE dataset.

The influence of age on nuclei volume

For the ADNI dataset, partial correlations, controlling for ICV, revealed a multiple comparison corrected significant influence of age on almost all nuclei separately in each group and over both groups. Overall, volume declined with age (see **Table 6**).

For the ABIDE dataset, AB, CAT, PL and Co were significantly correlated with age over both groups, while controlling for ICV. This effect was also present in three of those nuclei in the ASD group (AB, CAT, Co), but no analyses passed multiple comparison correction in the control group (see **Table 6**).

There were no significant differences in the relationship between nuclei volume and age between the groups for either dataset (after correction for multiple comparisons; see **Table 6**).

DISCUSSION

Here, we show that we can visualize the boundaries of 9 amygdala nuclei using high resolution *ex vivo* MRI data. The nuclei were consistent across cases and raters. Manual labeling of these nuclei in *ex vivo* MRI data served as a basis to construct a statistical atlas of the amygdala at the nucleus level. This amygdala atlas was applied to *in vivo* MRI data in two publicly available datasets (ADNI and ABIDE). We thus determined whether the atlas could be used to segment the amygdala nuclei in standard resolution T1 data of varying MR contrast. In addition this application showed how well the resulting segmentations could discriminate between Alzheimer's disease participants vs. control participants (ADNI dataset), and individuals with autism and age-matched controls (ABIDE dataset). We plan to incorporate the *ex vivo* amygdala nuclei atlas as part of the publicly available FreeSurfer software in the next release, thus opening up numerous multimodal applications in typical and atypical populations and allowing researchers to explore the amygdala nuclei's function and structure with greater specificity than previously possible in neuroimaging.

The amygdala is an important structure for animal and human cognition and serves as a crucial hub for cortical, subcortical and limbic connections throughout the brain. Extensive research from the animal literature and post-mortem human studies show that the amygdala is composed of several neuronal subpopulations (Freese and Amaral, 2005, 2006, 2009; Alheid, 2003; Price et al., 1987; Aggleton, 2000; Gloor, 1972, 1978, 1997; McDonald, 1998; LeDoux, 1998), however the precise functions thereof remain yet to be defined in detail in humans.

Currently, no *in vivo* parcellation methods available allow for an automated segmentation of nine amygdala nuclei derived from underlying anatomy and within standard imaging protocols. Previous studies have used *in vivo* imaging where the resolution did not reveal numerous amygdala nuclei, parcellating the amygdala into 2-4 regions (e.g. Entis et al., 2012; Solano-Castiella et al. 2011). T1 weighted scans of standard resolution (0.75-1 mm voxel size) do not provide sufficient overall image contrast for the human eye to distinguish the amygdala nuclei. Amygdala segmentation from diffusion weighted imaging data is possible, but again, the nuclei are usually grouped into larger subregions (e.g. basolateral instead of basal and lateral nuclei separately; Saygin et al. 2011, Bach et al. 2011). Compared with approaches that aim for estimating probability maps in reference spaces (e.g., Amunts et al., 2005, Solano-Castiella et al., 2010, 2011, Tyszka et al., 2016), the presented approach takes individual underlying anatomy into account, thus providing greater spatial sensitivity.

The present study offers four main innovations and advantages over previous work: i) higher (100-150μm isotropic) resolution and because of this high resolution ii) the largest number of amygdala nuclei resolved, labeled, and then modeled (nine nuclei), iii) an atlas based on n=10 cases thus allowing variability to be modeled across cases, and iv) the use of a recent modeling technique (Iglesias et al. 2015) that enables the inclusion in the atlas of large amounts of readily available *in vivo* segmentations of the whole amygdala. In particular, the generative nature of this model makes it agnostic to details of the imaging contrast, and hence permits the use of ultra-high-resolution, *ex vivo* training data. This is in contrast with techniques that require intensity matching between training and test data, which essentially forces the training data to be *in vivo* and thus of significantly lower resolution and overall image contrast.

In conventional probabilistic atlases defined in the space of a template brain, the usual procedure is to register the grayscale template to a test scan, and use the resulting transform to propagate the label probabilities. In contrast, for atlases based on generative models like the one presented here, the registered atlas only represents a prior, which is combined with a subject-specific likelihood model (agnostic to MRI contrast) to produce a posterior distribution that yields the segmentation in that individual's native space. Therefore, the method is adaptive to any MRI contrast, and since the prior and likelihood

inform each other's updates, it can generate sharper segmentations based on the posterior rather than the prior alone.

The atlas introduced in the current paper is easily applicable to standard anatomical MR data and will be implemented in a future release of the FreeSurfer software. FreeSurfer is an open source and widely used brain data analysis software and we intend that the new atlas will significantly impact future studies in advancing our understanding of amygdala nuclei function in the human brain.

Compared with approaches that aim for estimating probability maps in reference spaces (e.g., Amunts et al., 2005, Solano-Castiella et al., 2010, 2011), the present approach takes individual underlying anatomy into account, thus providing greater spatial sensitivity. While these previous studies contributed slightly larger sample sizes, the present study harnesses recent advancements in probabilistic modeling (i.e. Bayesian modeling) to take into account individual anatomy in order to parcellate the nuclei of the amygdala. Future work may also quantitatively compare the present atlas to other atlases that segment groups of nuclei (rather than individual nuclei as in the present atlas) in order to decide on the best atlases for different research questions. It is important to note that such comparisons of different atlases will depend heavily on the quality of the registration, which can be greatly compromised by differences in image contrast between the input atlases. When the registration of these template-space atlases is poor, it leads to poor overlap between the subregions segmented by different atlases and in different spaces (i.e. native subject space vs. template/group space).

Here, we also show that the new information about amygdala nuclei derived by the atlas can reliably distinguish between pathologic and normal anatomy in two separate populations, with higher accuracy than the volume of the whole amygdala. For the ADNI comparison, we expected gross anatomical differences between the pathological and the control group, given the neuropathology in AD and MCI (Merkesbery, 2010; Yilmazer-Hanke, 1998). In contrast, we expected rather small (if any) differences for the ABIDE group comparison. The exact structural neuropathology of the amygdala is less clear in ASD and there have been heterogeneous findings about the amygdala as a whole structure (e.g. Dziobek et al., 2010; Sparks et al., 2002). ASD is a very heterogeneous disorder clinically, and thus probably has a heterogeneous and complicated underlying

biology. Crucially, the accuracies of group discriminations based on the new atlas outperformed previously available results based on the amygdala as a whole in both the ADNI dataset ($p = 1.1\times10^{-5}$) and the ABIDE dataset ($p = .008$). Although the accuracy in the ABIDE comparison is not very high overall, using the current atlas still results in a significant improvement of discrimination accuracy. Moreover, the finding that the ASD discrimination is lower in accuracy than the Alzheimer's discrimination is another important aspect of the present analysis. Whereas the ADNI comparison extends previous findings of volumetric differences in the amygdala between AD and age-matched control participants, the findings in ASD show the potential for future research applications: using information about anatomical changes in amygdala nuclei in relation to behavioral characteristics can be more sensitive than the amygdala as one homogenous structure. We anticipate that this new tool will support neuroimaging researchers to find replicable and robust differences between ASD and controls with greater accuracy than possible based on the whole amygdala. We provide these results as an example of one potential application to a clinical population with less clear neuropathology and encourage other researchers to test and extend this finding in different ASD and other psychiatric samples respectively. For example, the estimation and definition of individual amygdala nuclei (or grouped into other meaningful subdivisions e.g. basolateral/centromedial complex) can serve as regions of interest for task-related functional MRI studies; and/or as seed regions for connectivity analyses in task-free fMRI data.

The additional exploratory analyses on the influence of hemisphere, sex, and age on nucleus volume in the *in vivo* datasets emphasizes the importance of careful matching of groups with respect to age and sex when comparing different groups (especially in older adult samples). Previous studies that have segmented the amygdala as a whole are discrepant, with some groups reporting larger right amygdala (e.g. Pedraza et al. 2002, Bernesconi N et al., 2003) or no asymmetry (Goncalves-Pereira PM et al., 2006). The absence of a hemispheric difference in the two datasets in the present study is informative and has implications for further investigations (e.g., volumes can be averaged), but in light of previous literature, future studies should carefully check for hemispheric differences. Another potential application could be to study volumetric and connectivity changes across development. For example, a recent study revealed that the connectivity

of most amygdala subregions (central, basal, and lateral) continued to change between the ages of 5-30 (Saygin et al. 2015). However, these regions were identified based on a connectivity atlas (Saygin et al. 2011) and the study was limited to coarser subdivisions of the amygdala. The structural (e.g. volume or connectivity) and functional developmental trajectory of the nine nuclei identified here remains to be explored.

Some limitations of the present study are that we did not label subdivisions of the amygdala nuclei (i.e. the ventral, dorsal, medial, lateral parts of each nucleus) and we did not label the "extended amygdala" in the ten cases. The medial and central nuclei are sometimes included as part of the extended amygdala, so perhaps we included part of the extended amygdala; others think such nuclei are part of the basal ganglia because of their inhibitory nature (Heimer & van Hoesen 2006; Olmos & Heimer, 1999; Cassell 1999, Swanson et al., 1998). It is also worth noting that different groups distinguish the central nucleus' boundaries quite differently; this discrepancy may be due to the two nuclei of the central nucleus (dorsomedial and ventrolateral), which are stained quite differently (described in Gloor (1978) on page 624). Amunts and colleagues included a larger central nucleus parcellation likely to underscore the extended amygdala concept in this region (Amunts et al., 2005). Here we predominantly label the medial portion of the central nucleus (i.e. dorsomedial), which would agree with many definitions of the central nucleus (Freese & Amaral, 2009) rather than the lateral portion of the central nucleus (i.e. ventrolateral), where some myelinated fibers separate it from the main mass of the amygdala. These differences between nucleic definitions among different groups will also be important to consider when directly comparing different atlases. The diversity of the amygdala nuclei with respect to its structure, function and connectivity - has been long debated. Future studies can clarify the roles of the different nuclei as part of the amygdala, ventral pallidum, basal forebrain, or hippocampus.

CONCLUSION

In this paper, we visualized nine amygdala nuclei boundaries (anterior amygdaloid area, cortico-amygdaloid transition areas; basal, lateral, accessory basal, central, cortical medial, paralaminar nuclei) using ultra-high-resolution *ex vivo* imaging. The nuclei were consistent across cases and raters, and the resulting atlas will be distributed in the

FreeSurfer software. The amygdala nuclei atlas was applied to two datasets, demonstrating a higher discriminability of Alzheimer's disease and Autism Spectrum Disorder than previously possible with amygdala segmentation methods. This amygdala atlas will provide neuroimaging researchers with the ability to test nucleus' function *in vivo* with greater spatial specificity in the human brain.

Acknowledgements

This work was supported by the PHS grant DA023427 and NICHD/NIH grant F32HD079169 (Z.M.S); Feodor Lynen Postdoctoral Fellowship of the Alexander von Humboldt Foundation (D.K.); R21(MH106796), R21 (AG046657) and K01AG28521 (J.C.A.), the National Cancer Institute (1K25CA181632-01) as well as the Genentech Foundation (M.R.); the European Union's Horizon 2020 Marie Sklodowska-Curie grant agreement No 654911 (project "THALAMODEL") and ERC Starting Grant agreement No 677697 (project "BUNGEE-TOOLS"); and the Spanish Ministry of Economy and Competitiveness (MINECO) reference TEC2014- 51882-P (J.E.I.); and the NVIDIA hardware award (M.R. and J.E.I.). Further support for this research was provided in part by the National Institute for Biomedical Imaging and Bioengineering (P41EB015896, R01EB006758, R21EB018907, R01EB019956, R01-EB013565), the National Institute on Aging (5R01AG008122, R01AG016495), the National Institute of Diabetes and Digestive and Kidney Diseases (1-R21-DK-108277-01), the National Institute for Neurological Disorders and Stroke (R01NS0525851, R21NS072652, R01NS070963, R01NS083534, 5U01NS086625), the Massachusetts ADRC (P50AG005134) and was made possible by the resources provided by Shared Instrumentation Grants 1S10RR023401, 1S10RR019307, and 1S10RR023043. Additional support was provided by the NIH Blueprint for Neuroscience Research (5U01-MH093765), part of the multi-institutional Human Connectome Project. In addition, BF has a financial interest in CorticoMetrics, a company whose medical pursuits focus on brain imaging and measurement technologies. BF's interests were reviewed and are managed by Massachusetts General Hospital and Partners HealthCare in accordance with their conflict of interest policies.

The collection and sharing of the ADNI MRI data used in the evaluation was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

References

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*(7021), 68-72.

Alheid, G. F. (2003). Extended amygdala and basal forebrain. *Annals of the New York Academy of Sciences*, *985*(1), 185-205.

Amunts, K., Kedo, O., Kindler, M., Pieperhoff, P., Mohlberg, H., Shah, N. J., ... & Zilles, K. (2005). Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. *Anatomy and embryology*, *210*(5-6), 343-352.

Ashburner, J., Andersson, J. L., & Friston, K. J. (2000). Image registration using a symmetric prior—in three dimensions. *Human brain mapping*, *9*(4), 212-225.

Augustinack, J. C., Huber, K. E., Stevens, A. A., Roy, M., Frosch, M. P., van der Kouwe, A. J., ... & Alzheimer's Disease Neuroimaging Initiative. (2013). Predicting the location of human perirhinal cortex, Brodmann's area 35, from MRI. *Neuroimage*, *64*, 32-42.

Bach, D. R., Behrens, T. E., Garrido, L., Weiskopf, N., & Dolan, R. J. (2011). Deep and superficial amygdala nuclei projections revealed in vivo by probabilistic tractography. *The Journal of Neuroscience*, *31*(2), 618-623.

Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., & Williams, S. C. R. (2000). The amygdala theory of autism. *Neuroscience & Biobehavioral Reviews*, *24*(3), 355-364.

Baxter, M. G., & Murray, E. A. (2002). The amygdala and reward. *Nature reviews neuroscience*, *3*(7), 563-573.

Bernasconi, N., Bernasconi, A., Caramanos, Z., Antel, S. B., Andermann, F., & Arnold, D. L. (2003). Mesial temporal damage in temporal lobe epilepsy: a volumetric MRI study of the hippocampus, amygdala and parahippocampal region. Brain, 126(2), 462-469.

Birbaumer, N., Grodd, W., Diedrich, O., Klose, U., Erb, M., Lotze, M., ... & Flor, H. (1998). fMRI reveals amygdala activation to human faces in social phobics. *Neuroreport*, *9*(6), 1223-1226.

Campeau, S., & Davis, M. (1995). Involvement of the central nucleus and basolateral complex of the amygdala in fear conditioning measured with fear-potentiated startle in rats trained concurrently with auditory and visual conditioned stimuli. *The Journal of Neuroscience*, *15*(3), 2301-2311.

Cassell, M. D., Freedman, L. J., & Shi, C. (1999). The intrinsic organization of the central extended amygdala. *Annals of the New York Academy of Sciences*,*877*(1), 217-241.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 837-845.

deOlmos, J. S. (2004) *Amygdala*. In: Paxinos. G., Mai, K.M.,editors. The Human Nervous System. Elsevier: 739-868.

Di Martino, A., Yan, C. G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., ... & Deen, B. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, *19*(6), 659-667.

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, *26*(3), 297-302.

Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science*, *298*(5596), 1191-1194.

Dolan, R. J. (2007). The human amygdala and orbital prefrontal cortex in behavioural regulation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *362*(1481), 787-799.

Dziobek, I., Bahnemann, M., Convit, A., & Heekeren, H. R. (2010). The role of the fusiform-amygdala system in the pathophysiology of autism. *Archives of general psychiatry*, *67*(4), 397-405.

Etkin, A., Klemenhagen, K. C., Dudman, J. T., Rogan, M. T., Hen, R., Kandel, E. R., & Hirsch, J. (2004). Individual differences in trait anxiety predict the response of the basolateral amygdala to unconsciously processed fearful faces.*Neuron*, *44*(6), 1043-1055.

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems.*Annals of eugenics*, *7*(2), 179-188.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341-355.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004b. Automatically parcellating the human cerebral cortex. Cereb Cortex 14, 11-22.

Freese, J. L., & Amaral, D. G. (2005). The organization of projections from the amygdala to visual cortical areas TE and V1 in the macaque monkey. *Journal of Comparative Neurology*, *486*(4), 295-317.

Freese, J. L., & Amaral, D. G. (2006). Synaptic organization of projections from the amygdala to visual cortical areas TE and V1 in the macaque monkey. *Journal of Comparative Neurology*, *496*(5), 655-667.

Freese, J. L., & Amaral, D. G. (2009). *Neuroanatomy of the primate amygdala*. In: Whalen P. J. , Phelps E.A.,editors. The Human Amygdala. New York: The Guilford Press.

Fujishiro, H., Tsuboi, Y., Lin, W. L., Uchikado, H., & Dickson, D. W. (2008). Co-localization of tau and α-synuclein in the olfactory bulb in Alzheimer's disease with amygdala Lewy bodies. *Acta neuropathologica*, *116*(1), 17-24.

Gloor, P. (1972). Temporal lobe epilepsy: its possible contribution to the understanding of the functional significance of the amygdala and of its interaction with neocortical-temporal mechanisms. In: Eleftherion, B.E., editor. The Neurobiology of the Amygdala. New York: Plenum Press; 1972.

Gloor, P. (1978). Inputs and outputs of the amygdala: what the amygdala is trying to tell the rest of the brain. In: LIVINGSTON, KE.; HORNIKIEWICZ, O., editors. Limbic Mechanisms The Continuing Evolution of the Limbic System Concept. New York and London: Plenum Press; 1978.

Gloor, P. (1997). The Temporal Lobe and Limbic System. New York, New York: Oxford University Press, Inc.

Goncalves-Pereira, P. M., Oliveira, E., & Insausti, R. (2005). Quantitative volumetric analysis of the hippocampus, amygdala and entorhinal cortex: normative database for the adult Portuguese population. *Revista de neurologia*, 42(12), 713-722.

Heimer, L., & Van Hoesen, G. W. (2006). The limbic lobe and its output channels: implications for emotional functions and adaptive behavior.*Neuroscience & Biobehavioral Reviews*, *30*(2), 126-147.

Hortensius, R., Terburg, D., Morgan, B., Stein, D. J., van Honk, J., & de Gelder, B. (2016). The role of the basolateral amygdala in the perception of faces in natural contexts. *Phil. Trans. R. Soc. B*, *371*(1693), 20150376.

Iglesias, J. E., Augustinack, J. C., Nguyen, K., Player, C. M., Player, A., Wright, M., ... & Van Leemput, K. (2015). A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. *NeuroImage*, *115*, 117-137.

Kalin, N. H., Shelton, S. E., & Davidson, R. J. (2004). The role of the central nucleus of the amygdala in mediating fear and anxiety in the primate. *The Journal of neuroscience*, *24*(24), 5506-5515.

Kliemann, D., Dziobek, I., Hatri, A., Baudewig, J., & Heekeren, H. R. (2012). The role of the amygdala in atypical gaze on emotional faces in autism spectrum disorders. *The Journal of Neuroscience*, *32*(28), 9469-9476.

Kotzbauer, P. T., Trojanowski, J. Q., & Lee, V. M. Y. (2001). Lewy body pathology in Alzheimer's disease. *Journal of Molecular Neuroscience*, *17*(2), 225-232.

LeDoux, J. (1998). Fear and the brain: where have we been, and where are we going?. *Biological psychiatry*, *44*(12), 1229-1238.

Leonard, J. A., Mackey, A. P., Finn, A. S., & Gabrieli, J. D. (2015). Differential effects of socioeconomic status on working and procedural memory systems.*Frontiers in human neuroscience*, *9*.

Markesbery, W. R. (2010). Neuropathologic alterations in mild cognitive impairment: a review. *Journal of Alzheimer's Disease*, *19*(1), 221-228.

McDonald, A. J. (1998). Cortical pathways to the mammalian amygdala. *Progress in neurobiology*, *55*(3), 257-332.

Olmos, J. S., & Heimer, L. (1999). The concepts of the ventral striatopallidal system and extended amygdala. *Annals of the New York Academy of Sciences*,*877*(1), 1-32.

Pedraza, O., Bowers, D. & Gilmore, R. (2004) Asymmetry of the hippocampus and amygdala in MRI volumetric measurements of normal adults. Journal of the International Neuropsychological Society, 10(5), pp. 664–678. doi: 10.1017/S1355617704105080.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception I: The neural basis of normal emotion perception. *Biological psychiatry*, *54*(5), 504-514.

Phillips, R. G., & LeDoux, J. E. (1992). Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behavioral neuroscience*, *106*(2), 274.

Price, J. L., Russchen, F. T., & Amaral, D. G. (1987). The limbic region. II. The amygdaloid complex. *Handbook of chemical neuroanatomy*, *5*(Part 1), 279-388.

Rauch, S. L., Shin, L. M., & Wright, C. I. (2003). Neuroimaging studies of amygdala function in anxiety disorders. *Annals of the New York Academy of Sciences*, *985*(1), 389-410.

Saygin, Z. M., Osher, D. E., Augustinack, J., Fischl, B., & Gabrieli, J. D. (2011). Connectivity-based segmentation of human amygdala nuclei using probabilistic tractography. *Neuroimage*, *56*(3), 1353-1361.

Saygin, Z. M., Osher, D. E., Koldewyn, K., Martin, R. E., Finn, A., Saxe, R., ... & Sheridan, M. (2015). Structural connectivity of the developing human amygdala.*PloS one*, *10*(4), e0125170.

Siegle, G. J., Steinhauer, S. R., Thase, M. E., Stenger, V. A., & Carter, C. S. (2002). Can't shake that feeling: event-related fMRI assessment of sustained amygdala activity in response to emotional information in depressed individuals.*Biological psychiatry*, *51*(9), 693-707.
Yilmazer-Hanke, D. M. (1998). Alzheimer's Disease. *Cells Tissues Organs*,*162*(1), 46-55.

Sims, K. S., & Williams, R. S. (1990). The human amygdaloid complex: a cytologic and histochemical atlas using Nissl, myelin, acetylcholinesterase and nicotinamide adenine dinucleotide phosphate diaphorase staining.*Neuroscience*, *36*(2), 449-472.

Solano-Castiella, E., Anwander, A., Lohmann, G., Weiss, M., Docherty, C., Geyer, S., ... & Turner, R. (2010). Diffusion tensor imaging segments the human amygdala in vivo. *Neuroimage*, *49*(4), 2958-2965.

Solano-Castiella E, Schäfer A, Reimer E, Türke E, Pröger T, Lohmann G, Trampel R, Turner R. (2011). Parcellation of human amygdala in vivo using ultra high field structural MRI. *Neuroimage*. 58(3), 741-8. doi: 10.1016/j.neuroimage.2011.06.047. Epub 2011 Jun 25. PMID: 21726652

Sparks, B. F., Friedman, S. D., Shaw, D. W., Aylward, E. H., Echelard, D., Artru, A. A., ... & Dager, S. R. (2002). Brain structural abnormalities in young children with autism spectrum disorder. *Neurology*, *59*(2), 184-192.

Swanson, L. W., & Petrovich, G. D. (1998). What is the amygdala?. *Trends in Neurosciences*, *21*(8), 323-331.

Tyszka, J. M. and Pauli, W. M. (2016). In vivo delineation of subdivisions of the human amygdaloid complex in a high-resolution group template. Hum. Brain Mapp.. doi:10.1002/hbm.23289.

Van Leemput, K., Bakkour, A., Benner, T., Wiggins, G., Wald, L. L., Augustinack, J., ... & Fischl, B. (2009). Automated segmentation of hippocampal subfields from ultra high resolution *in vivo* MRI. *Hippocampus*,*19*(6), 549-557.

Van Leemput, K. (2009b). Encoding probabilistic brain atlases using Bayesian inference. *Medical Imaging, IEEE Transactions on*, *28*(6), 822-837.

Figure legends

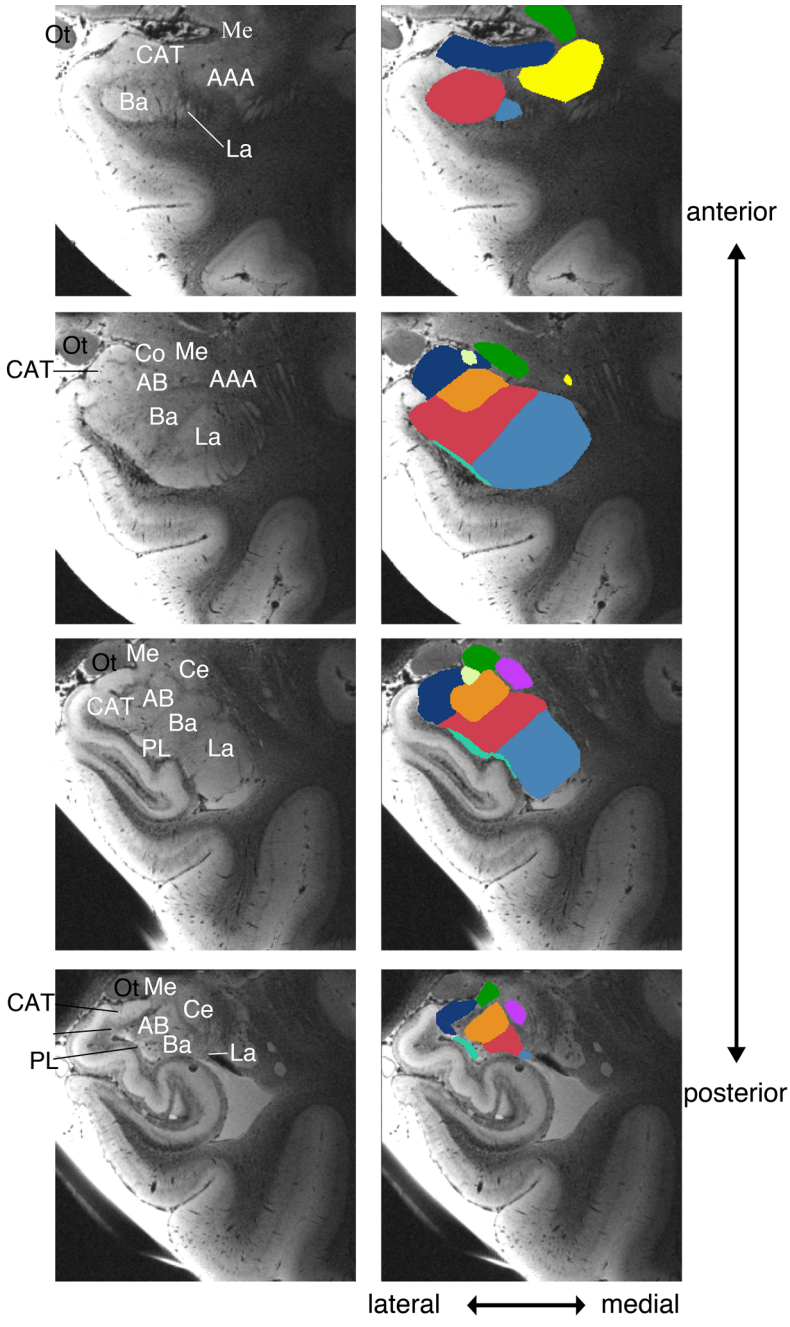
Figure 1. Coronal images from MRI of example *ex vivo* (case 7). The boundaries of nine amygdala nuclei were clearly visible on the left column and were used to hand-label the nuclei. Resulting nuclei labels illustrated on the right column. Slices extend from anterior to posterior amygdala (from top to bottom panels). La: lateral; Ba: basal; AB: accessory basal; Ce: central; Me: medial; Co: cortical; CAT: Cortico-amygdaloid Transition Area; AAA: Anterior Amygdala Area; PL: paralaminar nucleus; Ot: optic tract (as landmark).

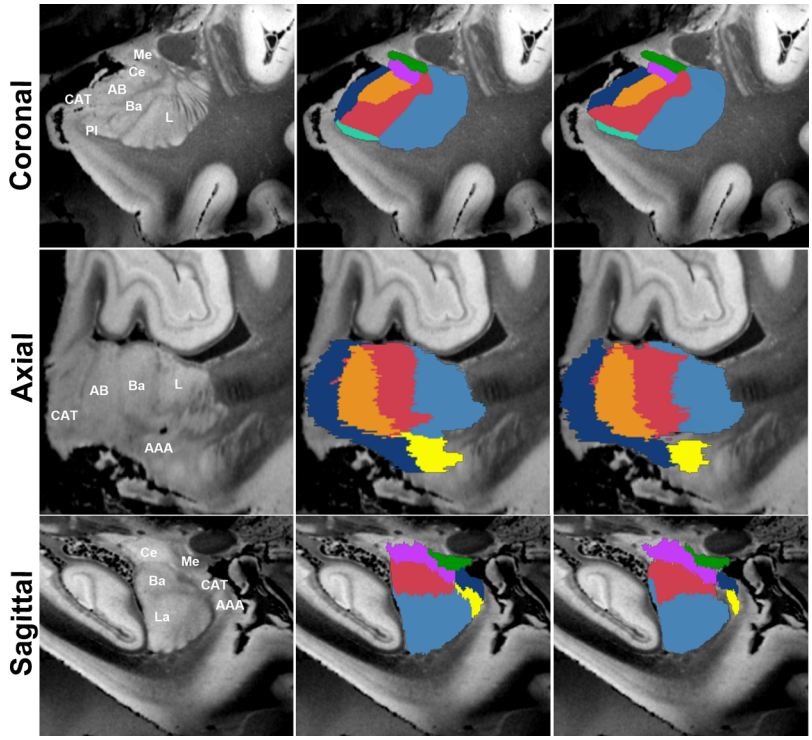
Figure 2. Inter-rater comparison of nucleus labels (case 1). Another example *ex vivo* case depicting the MRI contrast without any labels (left column) and with the manually-labeled nuclei produced by the two raters (middle and right columns). The location and spatial extent of the nuclei were similar between the two independent raters. Labels were based mainly on boundaries visible on coronal slices, but the two other orientations (axial and sagittal) were especially useful for checking boundaries of nuclei that were elongated in those orientations such as Co, CAT, Ce, and Me nuclei.

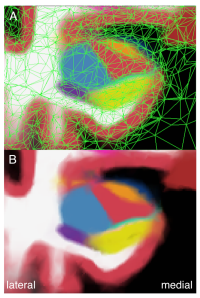
Figure 3. Coronal section of probabilistic atlas, with (A) and without (B) tetrahedral mesh superimposed. The color of each voxel is a combination of the colors of the different labels, weighted by the corresponding probabilities at each location. Different colors represent specific nuclei: green: Me, dark blue: CAT, orange: AB, red: Ba, purple: Ce off-white: Co yellow: AAA, light blue: LA, turquoise: PL.

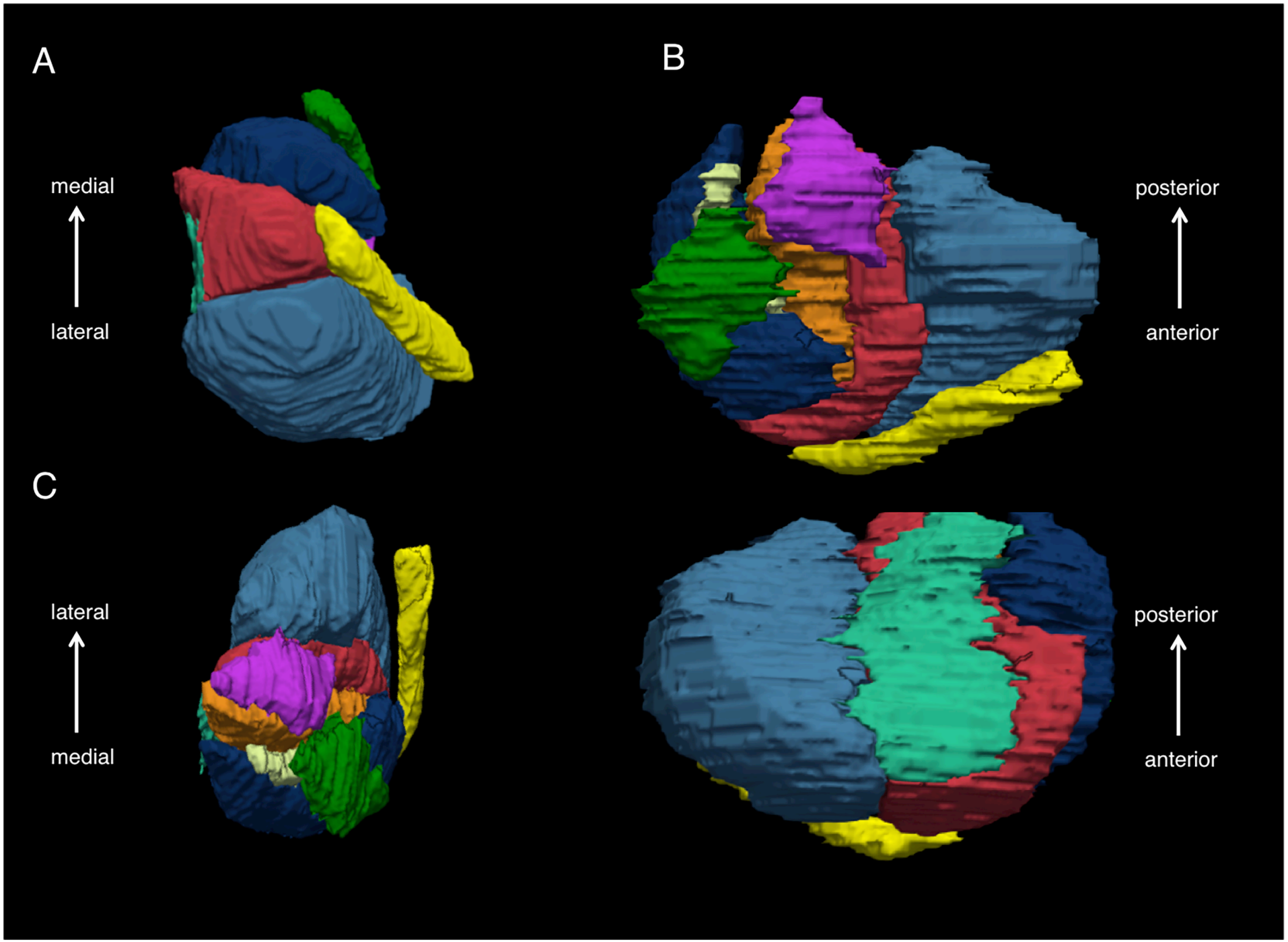
Figure 4. 3-Dimensional rendering of manual segmentation based on MRI in one *ex vivo* case. (A) anterior, (B) medial-lateral, (C) posterior , (D) coronal view. Different colors represent specific nuclei: green: Me, dark blue: CAT, orange: AB, red: Ba, purple: Ce off-white: Co yellow: AAA, light blue: La, turquoise: PL. For display purposes label boundaries are smoothed (5).

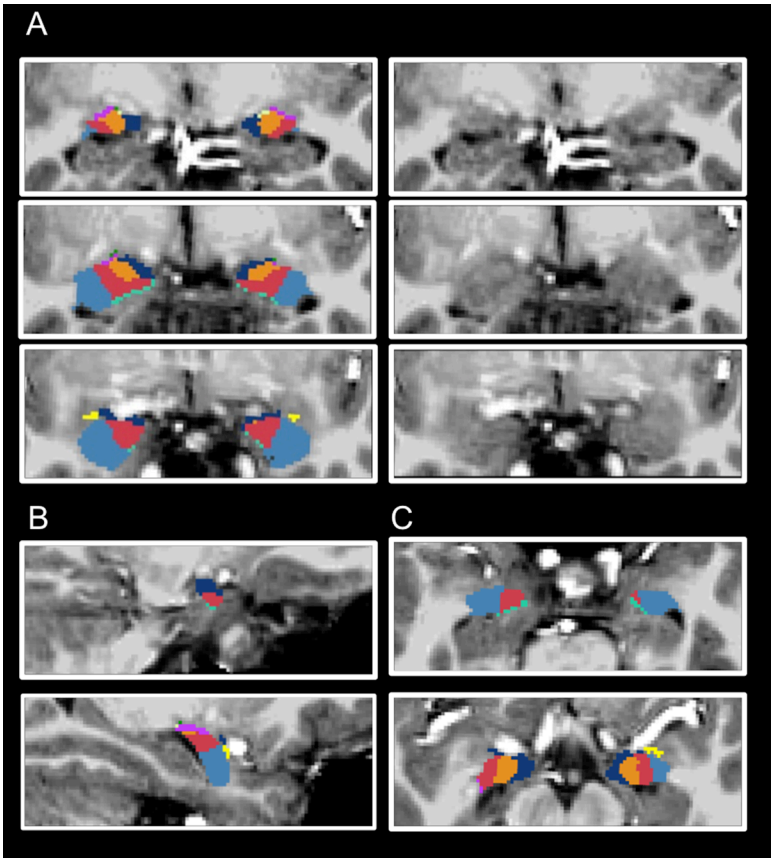
Figure 5. *In vivo* segmentations of amygdala nuclei overlaid on standard T1-weighted anatomical MR image (from ABIDE dataset). (A) Coronal, (B) sagittal, and (C) axial views. Panel A illustrates the MR image without any nuclei in order to visualize contrast quality. Different colors represent specific nuclei: green: Me, dark blue: CAT, orange: AB, red: Ba, purple: Ce, off-white: Co, yellow: AAA, light blue: La.


Ot
Me
CAT
AAA
Ba
La
anterior
Ot
CAT
Co
Me
AB
AAA
Ba
La
Ot
Me
Ce
CAT
AB
Ba
PL
La
CAT
Ot
Me
Ce
AB
Ba
La
PL
posterior
lateral
medial


Coronal
Axial
Sagittal
Me
Ce
AB
CAT
Ba
L
PI
AAA
La


A
B
lateral
medial


A
medial
lateral
B
posterior
anterior
C
lateral
medial
posterior
anterior


A
B
C

<u>Table 1:</u> Basic demographics and diagnostic information about brain samples used in this study. *Abbreviations*: AD, Alzheimer's disease; h, hours, m, male; f, female; *PMI*, post-mortem interval; *n/a*, data not available

Case #	Sex	Age	Laterality	Isotropic Resolution (μm)	Clinical Diagnosis	Neuropathology Diagnosis	PMI
1	n/a	n/a	left	150	control	control	< 24h
2	m	60	right	100	control	control	< 24h
3	f	86	left	100	mild AD	mild AD	18h
4	m	68	right	100	control	control	< 24h
5	m	n/a	left	120	control	control	< 24h
6	f	83	left	120	control	control	6h
7	m	63	left	120	control	control	< 24h
8	m	60	right	100	control	control	14h
9	m	68	right	100	control	control	<24h
10	m	58	right	100	control	control	<24h

Table 2. Overview of anatomical boundaries and landmarks for the manual labeling protocol.

Structure	***Abbreviation***	***Definition***
Anterior Amygdala Area	*AAA (yellow)*	The AAA represents the anterior end of the amygdala. AAA borders CAT anteriorly and laterally and has a concave crescent shape. In its most posterior and lateral position, AAA detaches from the rest of the amygdala and extends until striatal tissue becomes visible. AAA appears as a bright band anteriorly, similar to striatal tissue but AAA is more medial.
Cortico-amygdaloid Transition Area	*CAT (dark blue)*	The CAT represents the medial border of the amygdala. Laterally CAT borders AAA, AB, Ba, PL and Ce along its anterior-posterior extent. The posterior portion of CAT is inferior to the medial nucleus. CAT's ventral border merges into the hippocampal-amygdala transition area (HATA) posteriorly. Occasionally, the CAT showed poor contrast at its anterior borders.
Lateral Nucleus	*La (blue)*	In the anterior portion of the amygdala, the La is typically the first nucleus to appear. Scrolling anterior-posterior in the coronal plane, the La transforms from a circular/oval shape into a wedge or triangular shape. The La's medial border remains next to the Ba along the entire amygdala. The anterior La borders AAA, rostrally and laterally. The La continues laterally and dorsally until the posterior end of the amygdala. La is by far the largest nucleus of the amygdala, and reveals excellent contrast in all cases.
Basal Nucleus	*Ba (red)*	The anterior appearance of the Ba follows its lateral neighboring nuclei (La) and borders La throughout the amygdala. When viewed in coronal plane, Ba is circular anteriorly, then progresses into an L-shape midway, and ends circular.
Paralaminar Nucleus	*PL (turquoise)*	The PL is a small, light band that is inferior to Ba, lateral to CAT, and ventro-medial to part of the La. PL borders Ba and La and remains until the last few slices while transitioning more medially towards the CAT and AB.
Accessory Basal	*AB (orange)*	From anterior to posterior coronal slices, the AB emerges medially from/within the Ba in a circle that transforms into an oval shape. Dorsally, it forms an obtuse angle with Ba. Medially, the AB borders CAT, while its dorsal portion borders Ce in most of our cases.
Medial	*Me (green)*	The Me emerges near the optical tract and can be visible along most of the anterior-posterior extent of the amygdala. The Me covers most of the lateral-dorsal boundary of CAT. This nucleus is the most variable in shape, being either elongated and slim or more circular in coronal view. The axial view is useful in verifying the borders of this nucleus.
Central	*Ce (purple)*	The Ce appears circular and dorsal to AB and is between CAT medially and Ba laterally. For about half the cases, the Ce remains a circular shape, and for the other half of

		the cases, it becomes progressively more oval. The Ce appears brighter than its surrounding tissue. The axial view is useful in verifying the borders of this nucleus.
Cortical	*Co* *(off white)*	The Co emerges as a small circular nucleus, dorsally to CAT. The AB borders Co laterally. Overall, the Co was the smallest nucleus in size and contains the fewest number of slices labeled in our atlas.

Table 3. Mean volume of *ex vivo* nuclei across all cases used to create the atlas (mean mm^3 +/ se)

La	453.5 ± 31.4
Ba	300.9 ± 19.2
Ce	32.5 ± 7
Me	21.8 ± 5.6
Co	16.4 ± 3
AB	171.6 ± 16.9
CAT	174.8 ± 17.3
AAA	39.8 ± 7.9
PL	31.9 ± 6.4

Table 4. Accuracy and area under the curve results for discriminating AD vs. controls in ADNI dataset.

Volumes used as input	**Accuracy at elbow**	**AUC**	**p-value AD vs. controls**
Volume of whole amygdala from main FreeSurfer stream ("aseg", v5.1)	74.94%	0.844	5.68×10^{-31}
Volume of whole amygdala, (adding together the volumes of all nuclei, estimated with the new atlas)	81.46%	0.898	7.65×10^{-41}
Volumes of all 9 amygdala nuclei estimated with the new atlas, used simultaneously with LDA	84.07%	0.915	2.80×10^{-44}

Table 5. Accuracy and area under the curve for discriminating ASD vs. controls in ABIDE dataset.

Volumes used as input	**Accuracy at elbow**	**AUC**	**p-value ASD vs. controls**
Volume of whole amygdala from main FreeSurfer stream ("aseg", v5.1)	54.05%	0.4494	0.1605
Volume of whole amygdala, (adding together the volumes of all nuclei, estimated with the new atlas)	55.21%	0.4367	0.0544
Volumes of all 9 amygdala nuclei estimated with the new atlas, used simultaneously with LDA	59.46%	0.5902	0.012

<u>Table 6.</u> Correlation of age and nuclei volume per *in vivo* dataset (ADNI, ABIDE) for all subjects, disease group, control group and comparing correlations between groups. Abbreviations: Ncl., nucleus; p, significance value; r, Pearson's correlation coefficient.

		all subjects		*disease group*		*control group*		*group differences*	
in vivo	*Ncl.*	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*Fisher's z*
ADNI	AAA	**$4.9x10^{-5}$**	-.206	**$5.7x10^{-5}$**	-.306	**$2.3x10^{-5}$**	-.285	.826	-.22
	AB	**$5x10^{-6}$**	-.232	**$8.8x10^{-5}$**	-.299	**$2.5x10^{-11}$**	-.436	.103	1.63
	Ba	**$4.6x10^{-5}$**	-2.08	**$2.1x10^{-5}$**	-.323	**$7.5x10^{-7}$**	-.331	.928	.09
	CAT	**$3.9x10^{-4}$**	-.181	.01	-.198	**$1.7x10^{-7}$**	-.348	.116	1.57
	Ce	**$1.3x10^{-5}$**	-.221	**$3.3x10^{-5}$**	-.316	**$4.7x10^{-8}$**	-.363	.61	.51
	Co	**$4x10^{-6}$**	-.232	**$4.6x10^{-4}$**	-.268	**$7.9x10^{-9}$**	-.382	.219	1.23
	La	**$8.2x10^{-7}$**	-.249	**$3.3x10^{-5}$**	-.315	**$1.5x10^{-7}$**	-.350	.703	.38
	Me	**.002**	-.157	.211	-.097	**$7x10^{-6}$**	-.302	.039	2.07
	PL	**$8.6x10^{-4}$**	-.170	**$3.0x10^{-5}$**	-.317	.007	-.183	.168	-1.38
ABIDE	AAA	.745	.02	.879	.014	.813	.021	.745	.020
	AB	**.001**	-.211	**.002**	-.238	.116	-.139	**.001**	-2.11
	Ba	.019	-.146	..056	-.169	.160	-.124	.019	-.146
	CAT	**$3.6x10^{-4}$**	-.220	**.003**	-.257	.035	-.186	**$3.6x10^{-4}$**	-.220
	Ce	.097	-.104	.012	-.221	.457	.066	.097	-1.04
	Co	**.003**	-.186	**.002**	-.267	.340	-.085	**.003**	-1.86
	La	.828	-.014	.541	-.054	.604	.046	.828	-.014
	Me	.606	-.032	.103	-.145	.119	.138	.606	-.032
	PL	**.003**	-.183	.020	-.205	.070	-.160	**.003**	-.183