SAHLGRENSKA ACADEMY

ATLAS DATABASES OF HUMAN BRAIN ANATOMY: QUALITATIVE AND QUANTITATIVE COMPARISON

M.Sc. Thesis

Amanda Philipsson Franzén

Essay/Thesis:	30 hp
Program:	Medical physics
Level:	Second Cycle
Semester/year:	Spring 2020
Supervisor:	Rolf A. Heckemann
Examiner:	Magnus Båth

Abstract

Essay/Thesis:	30 hp
Program:	Medical physics
Level:	Second Cycle
Semester/year:	Spring 2020
Supervisor:	Rolf A. Heckemann
Examiner:	Magnus Båth
Keyword:	Brain atlas concordance problem, Quantitative comparison, Qualitative comparison, $Hammers_{mith}$, 2012 MICCAI Multi-Atlas Labelling Challenge Data.

Purpose: The purpose of this study was to quantitatively and qualitatively compare the two brain atlases databases, $Hammers_{mith}$ and 2012 MICCAI Multi-Atlas Labelling Challenge data, by developing a quantitative method.

Theory: Anatomical atlases of the human brain provide reference information about its structure. Researchers and practitioners use them for varied purposes such as automatic image segmentation, biomarker discovery, and identification of relationships between brain structure and function. There is no worldwide agreement on how to segment the human brain, which gives rise to difficulties and differences in the description of brain structures: the brain atlas concordance problem. Two widely used atlas databases are investigated in this study: the $Hammers_{mith}$ (HM) and the 2012 MICCAI Multi-Atlas Labelling Challenge Data (MGC). Both consist of T1-weigthed 3D magnetic resonance (MR) brain images of 30 study participants, with corresponding anatomical label sets.

Method: The study data consisted of 60 MR brain images (30 from each database) with 120 corresponding segmentations (30 manual and 30 automatically generated, times two databases). The automatic segmentations of the MGC images were based on the HM atlas, and the automatic segmentations of the HM images were based on the MGC atlas. The study was composed of two main parts, a qualitative comparison and a quantitative comparison. The quantitative comparison was developed during the study and was evaluated by juxtaposition with the qualitative results. The quantitative method included calculation of the most frequent coinciding regions, the Jaccard coefficient, and the volume ratio between corresponding regions from each database. The qualitative comparison was composed of predicting differences based on a comparison between the delineation protocols for a subset of regions, a visual analysis of overlaps and a global comparison of region names included in the protocols.

Result: The main difference between the protocols is that cortical regions only include the actual cortical grey matter in the MGC, whereas HM includes adjacent white matter as part of the region. 73 of the HM regions had matching region names with the MGC regions and 86 of the MGC regions had matching region names with the HM regions. The main differences between the defined regions from the databases were the subdivisions of regions and inclusion of different gyri. 40 and 43 of the HM regions had the matching MGC region as the most frequent coinciding region in the HM images respective the MGC images, and 76 of the MGC regions had the matching HM region as the most frequent coinciding region in both the HM and MGC images. Both atlas databases leave certain brain regions unclassified (assigned the background label). The Jaccard coefficient showed that the greatest overlap occurred between the regions that had matching names with regions in the other protocol. The HM regions generally had larger volumes compared to the corresponding MGC regions, although there were exceptions where MGC regions were almost twice the size of the corresponding HM regions. The quantitative comparison confirmed most of the predictions and revealed multiple additional overlaps and insights that could not be predicted just based on studying the protocols.

Conclusion: The two atlas databases differ systematically, reflecting the differences in purpose and priorities that guided the underlying manual segmentation procedures. The quantitative method developed in this project showed to be able to confirm the most important predictions and reveal additional insight that the qualitative analysis could not predict.

Table of content

1 Introduction

Anatomical atlases of the human brain provide reference information about its structure. Researchers and practitioners use them for varied purposes such as automatic image segmentation, biomarker discovery, and identification of relationships between brain structure and function. There are many different databases that contain such information, that integrate images from modern modalities and play an important role in clinical neuroscience.

1.1 The brain atlas concordance problem

The variety of available brain segmentation protocols reflects the diversity of purposes and motivations for constructing atlases. For this reason, there are, for example, atlases based on cytoarchitecture or landmarks with varying degrees of subdivision; functional and connectivity-based segmentations, and multi-modal segmentations (Nurbaya Yaakub et al., 2020). There is no worldwide agreement on how to segment the human brain, which gives rise to difficulties and differences in the description of brain structures: the brain atlas concordance problem. Some of the difficulties that have hampered the development of a standard delineation protocol are "… a dearth of identifiable landmarks, inter-subject variability, and imprecise or indeterminable structure–functional relationships…" (Bohland et al., 2009, p. 2). The inconsistency in neuroanatomical nomenclature has traditionally been viewed as to compose the problem and to include two key components. The first is that the same anatomical or functional brain structure has been referred to by multiple names and the second is that the same name has been used for different regions. Therefore, the problem has mainly been handled through the compilation of large lists of neuroanatomical region labels, attempts to build thesauri for relating these terms, and by developing machine-readable controlled vocabularies and ontologies. Although some atlases have similar region names, the corresponding regions do not necessarily coincide. Accordingly, comparing published results that uses regions from one anatomical atlas to regions of another atlas requires more than matching region names. A precise, quantitative understanding of the correspondence between the underlying anatomical divisions and a quantitative method to describe discrepancies would therefore be desirable (Bohland et al., 2009).

1.2 Hammers$_{mith}$ and the 2012 MICCAI Multi-Atlas Labelling Challenge Data

Two widely used atlas databases are investigated in this study: the Hammers$_{mith}$ (HM) and the 2012 MICCAI Multi-Atlas Labelling Challenge Data (MGC). Both consist of T1-weighted 3D magnetic resonance (MR) brain images of 30 study participants, with corresponding anatomical label sets. A label set consists of anatomical structures that are manually delineated according to specific protocols. These two databases were created independently of each other, though they still share some common characteristics. For example, the anatomical label sets have been generated by trained experts and the labels are integers that match with the MR images voxel-to-voxel, specifying the neuroanatomical region present at that voxel (Neuromorphometrics Inc., 2018). The main differences between these two databases lie in the delineation protocols.

1.3 Parcellation and segmentation

Parcellation and segmentation are terms that are commonly used in many articles discussing brain atlases. They are often used interchangeably. However, parcellation would best be reserved to refer to two-dimensional segmentations on surfaces, while segmentations are all kind of partitions into

separate parts or sections. Segmentations can be both cortical and subcortical while parcellations mostly can be done on cortical regions.

1.4 MAPER

Manually generated atlases can be used to automatically segment novel target MR brain images. One method to do this is called MAPER (multi-atlas propagation with enhanced registration) (Heckemann et al., 2010). The MAPER software automatically creates new atlases of new brain images using manually segmented atlases as reference and taking the overall brain structure into account during the registration (Nurbaya Yaakub et al., 2020). It is based on the principal idea of transferring knowledge from an atlas to a target image and uses multiple manually segmented atlases. This method was developed using the HM atlases, but other atlases can be used as well (Heckemann et al., 2006) (Heckemann et al., 2010). MAPER has been used to, for example, process the 996 baseline and screening images of Alzheimer's Disease Neuroimaging Initiative, for connectivity-based subsegmentation of the thalamus and to correct partial volume effects in opioid receptor PET (Heckemann et al., 2011).

1.5 Choice of atlas

Most analyses of brain images do not directly depend on anatomical segmentations. However, the choice of anatomical reference atlas could affect the way results are interpreted, reported and compared with other results. It is therefore important to know which atlases one uses and how it can influence the result. The number of available segmentation methods are increasing and the need for quantitative methods that capture the relationship between different protocols and enable mapping between them is thereby also increasing (Bohland et al., 2009).

1.6 Aims

The aim of this project was to provide a simple and reliable method to qualitatively and quantitively compare atlases or segmentations. The method should provide a way to predict the relationship between the protocols and enable mapping between them. This was done by using multiple easily available measures to quantitatively and qualitatively describe discrepancies between atlases which were applied to a pair of atlas databases and thereby showing how they differed quantitatively and providing a proof-of-concept in the process.

2 Method and Materials

This study consisted of both a qualitative and a quantitative comparison between the two brain atlas databases, HM and MGC. The qualitative comparison was mainly composed of a comparison between how the delineation of different regions were made, a visual analysis of overlaps and a global comparison of region names included in the protocols. The quantitative comparison included calculation of the most frequent coinciding region, the ratio between volumes and the Jaccard coefficient. The quantitative method that was developed was then evaluated using four regions and the qualitative comparisons of these. The study data consisted of 60 MR brain images (30 from each database) with 120 corresponding segmentations (30 manual and 30 automatically generated, times two databases). The automatic segmentations of the MGC images were based on the HM atlas, and the automatic segmentations of the HM images were based on the MGC atlas. Automated segmentation was done using the MAPER software. The HM images had voxels that were 0.94 x 0.94 x 0.94 mm while the MGC had images with voxels that were 1.0 x 1.0 x 1.0 mm. Each image had different matrix dimensions.

2.1 The HM database

The creation of the HM database was led by the neurologist Alexander Hammers, who has a great interest in epilepsy. This interest was reflected in that the early versions of the whole-brain atlases had finer subdivisions of the temporal lobe. The HM atlases have been created bit by bit, with new structures added gradually. The segmented regions have been carefully prepared using detailed, validated protocols. The first atlases were based on T1-weighted MR brain images from 20 healthy young adults and included 49 regions of interest. The anatomical structures were delineated by one investigator on each MRI before the next structure was initiated. Each structure to be delineated was assigned a unique voxel value which can be translated on screen as a greyscale intensity or to a colour. To make sure that there had been no development in the rendition of the protocol, each specific structure was re-examined after it had been delineated on all of the MRIs and on both hemispheres. In all cases where a general agreement was not reached, a neuroanatomically trained operator was consulted (Hammers et al., 2003). The segmentation protocol was later extended to 83 regions of interest when a proposed method for automatically creating anatomical atlases was developed. This method was applied to an exemplar data set of volumetric MR images from children at 2 years of age. 30 atlases created from 30 T1-weighted MR images from healthy adult volunteers was used in this study. The trained raters who manually performed the delineation were blinded to the demographic information. The manually used cursor in the Analyze AVW software was used to define the boundaries of the macroanatomical regions (Gousias et al., 2008). In Wild et al. (2017), the parietal lobe was divided into four regions, further extending the segmentation protocol into 87 regions. Around the same time the segmentation protocol was finally extended to its current 95 regions by subdividing the insula into six regions. The manual delineation of the insula was made using Rview v9, where the form, structure and surface of each region was examined in three orthogonal views of the MR image. The borders were delineated first to assign a region to all the grey matter voxels and then the rest of the white matter that wasn't assigned to a region was assigned to each subdivision (Faillenot et al., 2017). All of the 30 MR images were obtained on the 1.5 Tesla GE Signa Echospeed scanner at the National Society for epilepsy. These 30 MR images were obtained from 15 males and 15 females with the age range of 20–54, with a median age of 31 years, where 25 of the subjects were strongly right-handed (Gousias et al., 2008).

2.2 The MGC database

The MGC data consists of T1-weighted MR images from 30 subjects, with the age range of 18–96 years and all right-handed, from the Open Access Series of Imaging Studies database (Marcus et al., 2007). The protocol defines 207 labels, however only 138 cortical and subcortical labels occur in the data set. The manually delineated atlases were generated by Neuromorphometrics Inc. and was provided for use in the MICCAI 2012 Grand Challenge and Workshop on Multi-Atlas Labelling (Nurbaya Yaakub et al., 2020). Neuromorphometrics Inc. offers brain measurement services given raw MRI brain scans. The automated analyses are manually guided, inspected and certified by a neuroanatomical expert. Each structure to be delineated was assigned a unique voxel value here as well. Two segmentation protocols are used to precisely define the region borders. The "general segmentation" defined by the MGH center of Morphometric Analysis was the foundation for the first protocol. BrainCOLOR Cortical Parcellation Protocol was used to segment the cerebral cortex into regions defined by gyral and sulcal landmarks. The boundaries of each structure were located and then the voxels inside the region was assigned the label defining the region. The delineation was made by neuroanatomical technicians using Neuromorphometrics Inc.'s software NVM. Delineations were made on each MRI slice and any debatable neuroanatomy was evaluated by a neuroanatomist (Neuromorphometrics Inc., 2018).

2.3 Qualitative comparison

The qualitative comparison consisted of three parts: making predictions from the delineation protocols of between which regions, from the two atlases, overlap would occur as well as where geometrically these overlaps would occur, plotting each region and its overlap with other regions and visually analysing them and finally making a global comparison between the defined regions in each protocol based on their names.

2.3.1 Predictions from protocols

The qualitative comparison was initiated by comparing the two segmentation protocols for four selected regions. The four regions chosen were the hippocampus, the thalamus, the cingulate gyrus and the fusiform gyrus. The protocol comparison was performed by comparing which view the region was drawn in, comparing the boundary definitions in each direction separately and in its entirety, and then drawing conclusions about how the differences would show when the atlases were overlapped by each other. All of the protocols describing the HM segmentations of the four regions were found in Hammers et al. (2003). The MGC segmentation protocol for the hippocampus was found at Neuromorphometrics Inc. (2005), protocol for the thalamus at Neuromorphometrics Inc. (2005), and protocol for the cingulate gyrus and fusiform gyrus was found at Tourville et al. (2010). The nomenclature of the regions in the protocols differed slightly. The nomenclature that was used to compare the thalamus was "thalamus" in the HM protocol and "thalamus proper" in the MGC protocol. The nomenclature used for the cingulate gyrus was "gyrus cinguli" in the HM protocol and "cingulate gyrus" in the MGC protocol. The cingulate gyrus was divided into an anterior and a posterior part in the HM protocol and the MGC protocol divided the gyrus into an anterior, a middle, and a posterior part. The anterior part defined by the HM was compared to the anterior and middle part defined by the MGC protocol and the posterior part defined by the HM was compared to the posterior and middle part of the gyrus defined by the MGC. The anterior and posterior part defined by the HM was also compared to the anterior, middle and posterior part defined by the MGC when all of the parts were put together. Figures from a supplement to Gousias et al. (2008) (illustrated delineation protocols

obtained by request from the corresponding author) were used to compare the cingulate gyrus. The specific figures that were used were Fig (5051-3), Fig II-2 and Fig II-3 to find the boundary between the anterior and posterior part and to visualize the inferior border of the posterior part. The nomenclature for the fusiform gyrus in the HM protocol was “lateral occipitotemporal gyrus (fusiform gyrus)” and in the MGC protocol the corresponding region was divided into the “fusiform gyrus” and the “occipital fusiform gyrus”. Both parts of the MGC region were compared to the one corresponding region in the HM protocol. Every border that had a corresponding border in the other protocol was compared and predictions of how they related to each other and possible differences between them were written down. If the border in one protocol did not have a corresponding border it was not compared, though it was considered when predictions about the entire region were made.

2.3.2 Visualization

Evaluation of the predictions was partly done by visualizing the overlaps between the four regions and the regions they according to the predictions overlapped with. This was done using the *ortho_diff* function from the Neurobase package in R (Muschelli, 2020). Some of the overlaps that were predicted for each of the four regions were plotted and visualized in four different brain images. Each overlap was visualized in three orthogonal views (coronal, sagittal, and transversal) on the slice chosen by the plotting function. The plot showed the MGC atlas in the background in grey scale, and the investigated regions were compared to each other and plotted as voxels that had different colours depending on if the voxel were false positive, false negative, or true positive. False positive implied that the voxel belonged to the MGC region but not the HM region, false negative implied that the voxel belonged to the HM region but not the MGC region, and true positive implied that the voxel belonged to both the HM and MGC region.

2.3.3 Global comparison of the region names

A global comparison between defined regions in the MGC and the HM was carried out to investigate which extra regions the MGC defined and to determine if there were any regions in the HM that had no corresponding region in MGC. The comparison was made by comparing the names of the regions. Only the regions that had an obvious corresponding region based on its name were regarded to have a matching region in the other protocol. Subdivisions that had names which clearly matched a larger region in the other protocol were regarded as having a matching region, but regions which only seemed to be partly covered by a region were not regarded as a match. For example, the HM defined the insula anterior long and short gyrus, which were matched with the anterior insula defined by the MGC, but the MGC region anterior insula was not regarded as having a matching region. In other words, multiple subdivisions could be matched to one region in the other protocol, which made it possible for one protocol to have more matching regions than the other. The MGC defined 207 regions in their protocol, but only 138 occurred in the data set. These 138 regions can be found in Ourselin (2014). These 138 regions were compared with the 95 regions defined in the HM protocol.

2.4 Quantitative comparison

The first step in the quantitative comparison was to overlap the manually generated atlases with the automated generated atlases. The automated generated atlases corresponding to the MGC images using the HM atlas data were overlapped with the manually generated atlases of the MGC images, and the automated generated atlases using the MGC atlas data for the HM images were overlapped with the manual generated atlases of the HM images. To be able to analyse the overlapped atlases and determine which voxels corresponded to which combination of regions, the MGC labels were first

multiplied by 96, then the HM label set was added to the corresponding MGC label set for each brain image. By using modulo calculation, the combinations of regions were decoded. The quantitative comparison was mainly composed of three parts: determination of the most frequent coinciding combination of regions, calculation of the Jaccard coefficient and the comparison of volumes of each region. The MGC and HM images and their corresponding atlases were analysed separately. The comparison was carried out using RStudio.

2.4.1 Most frequent coinciding combinations

The most frequent coinciding overlap was first determined by adding all of the combined atlases that belonged to the respective database images and counting all of the voxels that belonged to each combination of regions. A table was generated that contained the rank of coinciding MGC regions by frequency for each HM region and vice versa, the number of voxels that belonged to each combination, and the names and numbers representing the HM and MGC regions each combination represented. From this table it was possible to determine, for example, which the most frequent MGC region was for every HM region and vice versa, and to determine the ratio between the number of voxels that belonged to the combination with the second and first most frequent coinciding region for each MGC region and each HM region. The ratio between the number of voxels that belonged to the combination with the third and first most frequent coinciding region was also calculated. This was also done for each combined atlas (i.e. for the 60 brain images the HM and MGC atlas were overlapped) individually and the mean, standard deviation and the coefficient of variance (CoV, standard deviation divided by the mean) of the 60 combined atlases (30 for each database) for each ratio for every combination was calculated and analysed. The number of combined atlases that had the most common most frequent coinciding region was also added to the table.

The ratios were analysed to understand the overlap and how much each region overlapped with the chosen region, to determine if the overlap was systematic or random as well as to understand why there might be some inconsistency between the images regarding which region was most frequently coinciding. The most frequently coinciding region was also compared to the global comparison to see which regions had the expected combination as the most frequently coinciding.

2.4.2 Jaccard Coefficient

The Jaccard coefficient was used to quantify the overlap between the manually delineated and automatically generated regions. This measurement was first mentioned as a coefficient of community when Paul Jaccard was investigating the distribution of the flora in the Swiss alps (Jaccard, 1912). The Jaccard coefficient is defined as the intersection divided by the union and is calculated as follows:

$$JC = \frac{A \cap B}{A \cup B}.$$

In this study, A and B represents the pair of HM and MGC regions, the intersection is the number of voxels that belongs to both regions, and the union is the total number of voxels that belong to either one of the regions or both (see Figure 1).

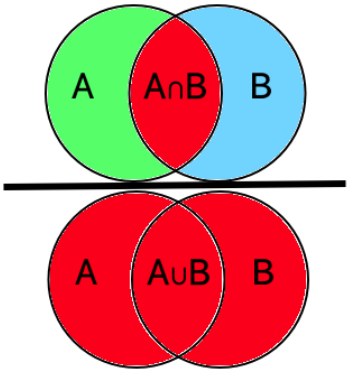

Figure 1: Schematic of the intersection and union used to calculate the Jaccard coefficient.

The coefficient was calculated for each combination of regions for both when all atlases were added to each other and for each combined atlas separately. The mean, standard deviation and CoV of the Jaccard coefficient for every combination was calculated between the 30 combined atlases belonging to each database images. The Jaccard coefficient was analysed by registering for which combination of regions the largest overlap occurred, investigating if there were a correlation between the region pairs, comparing the CoV between large and small overlaps, and determining whether there was any combination that had a deviating variation. The number of combined atlases the overlaps occurred in was also calculated.

2.4.3 Volumes

The volume of each region defined by the HM and the MGC was calculated by adding all of the voxels that belonged to each region. The ratio between the corresponding regions was calculated and used as a measurement to compare the region volumes between the databases. The ratio was calculated by:

$$r = \frac{HM\ region}{MGC\ region},$$

which meant that a value greater than one represents that the HM volume is larger than the MGC volume.

The mean, standard deviation and CoV of each region and ratio was calculated using the 30 combined atlases from each database to get an idea of how much they varied between the different brains. This was analysed by investigating whether there was a pattern between the CoV of the volume and the CoV of the volume ratio, whether there was a correlation between the volume and which hemisphere the regions belonged to and investigating the ratio for the regions that had a matching region.

2.5 Evaluation of the quantitative comparison

To evaluate if it was possible to predict the actual overlap between regions based only on the delineation protocols and to evaluate the quantitative method, a comparison between the qualitative predictions of the four regions and the quantitative calculated overlaps was performed. The evaluation consisted of analysing the Jaccard coefficient, the region volumes and ratios, the number of regions each region overlapped with and the most frequent coinciding regions of the four selected regions and

comparing these with the qualitative predictions. If there was an overlap that had not been predicted, the protocol for the investigated region as well as for regions nearby was reviewed again to see whether the overlap could be explained.

3 Results

3.1 Predictions from protocols

3.1.1 Hippocampus

The anterior border is approximately the same, since for the MGC the border is the amygdaloid nuclear complex, and it is stated that when the temporal horn of the lateral ventricle is visible in the slice, it is likely that the hippocampus is present. The HM defines the border as where the temporal horn loses it slit-like appearance, widens, and lies next to the hippocampus. They may differ somewhat since the amygdaloid nuclear complex and the temporal horn are not exactly in the same spot. The lateral border consists of the temporal horn of the lateral ventricle for the MGC and the lateral ventricle and white matter for the HM, which implies that these borders are similar. The MGC define the posterior border as under the pulvinar while the HM define the border as where the cella media, temporal horn and occipital horn fuse and exclude the hippocampal tail. This give rise to a difference between the borders in the posterior part. The medial border is defined as CSF for both of the protocols, so they are the same. The MGC inferior border will be more inferior than the HM border because the MGC defines the border as the white matter between the entorhinal cortex and posterior parahippocampal gyrus, and the HM defines it as the parahippocampal gyrus, uncal sulcus, interface of the prosubiculum and cornu ammonis and border between subiculum, and praesubiculum; sulcus hippocampalis. The MGC medial inferior border includes the subiculum, most of the presubiculum and about a quarter of the parasubiculum, while the HM defines the border as the uncal sulcus, interface of the prosubiculum and cornu ammonis, as well as the border between subiculum and presubiculum. Accordingly, the MGC includes all of the subiculum while HM excludes some of it, which will affect the overlap and volume.

The MGC volume will be a bit larger than the HM volume of the hippocampus since MGC defines it to extend more inferior and posterior, but otherwise the two structures overlap pretty well.

3.1.2 Thalamus

The anterior border will be the same, since both protocols define the border as the foramen of Monroe. Posteriorly the MGC extend the thalamus more posterior than the HM, since it overlaps the midbrain, and CSF is the border, while the HM defines the border as the slice where the pulvinar is first visible. The superior border will be the same since MGC's border is the transverse cerebral fissure, which is a fissure between the corpus callosum and the fornix above the thalamus and the roof of the 3rd ventricle below, and HM's borders are white matter / corpus callosum posteriorly. White matter and the lateral ventricle are defined as superior borders for both of the databases. In the anterior part of the superior border, the thalamus borders the caudate according to MGC and the stria terminalis / vena thalamostriata according to HM, which is marking a line of separation between the thalamus and the caudate nucleus. Since the MGC does not exactly define where the border go, it is a bit anterior superior of the HM border. The inferior border is slightly different in both the posterior and anterior part. This since the HM border is the cisterna ambiens while the MGC uses the hippocampus as the most posterior inferior border. In the anterior part, MGC defines the hypothalamic fissure as a border (dividing the thalamus and ventral diencephalon) while HM uses the anterior temporal lobe (both medial and lateral part), and therefore the border is CSF. Because of this, the HM inferior anterior border extend further than the MGC border, since the HM does not define a ventral diencephalon region. The lateral border is approximately the same, i.e. the internal capsule, but while the HM uses white matter of the temporal horn and the insula, the MGC uses the intensity contour function to mark

the lateral border. This implies that MGC has a tighter lateral border than HM. The medial border is also approximately the same, the databases define the border as CSF posterior and the third ventricle anterior.

The anterior, superior, inferior, lateral and medial border is approximately the same and the regions overlap each other well. The volume of the regions is approximately the same but the MGC region will be slightly larger.

3.1.3 Anterior cingulate gyrus

The anterior border for both of the databases is defined as the sulcus cinguli. The posterior border in MGC is defined for the sub- and supracallosal area separately: for the subcallosal area the border is the limit of the medial frontal cortex, and for the supracallosal area the border is the posterior limit of the genu of the corpus callosum (genu-post, plane marked by the posterior limit of the genu of the corpus callosum). The posterior border of the HM is instead made up by vertical lines drawn from the corpus callosum to the sulcus cinguli at the midpoint of the greatest extension of the corpus callosum and the corpus callosum inferiorly. These borders are quite different: the HM anterior part will extend further posteriorly than the MGC. The greatest extent of the corpus callosum is approximately in the middle of the corpus callosum. The superior border is the same for the MGC and the HM: both define the border as the cingulate sulcus, and if a double sulcus cinguli is present, the anterior one mark the border. The MGC also defines the subcallosal area border as the callosal sulcus, which cause some differences in the borders between the databases. The inferior border according to MGC is the callosal sulcus in the supracallosal area and the superior rostral sulcus / posterior projection from the posterior limit of the superior rostral sulcus in the subcallosal area. The HM defines the border as the most inferior slice on which the genu corporis callosi is uninterrupted throughout its width. The subcallosal inferior border is further inferior than the HM defined border, since the rostral sulcus is located more inferior than the most inferior slice of the genu corporis callosi. The lateral border is only defined by the HM. The middle part of the cingulate gyrus defined by the MGC will overlap both the anterior and posterior part of the cingulate gyrus defined by the HM, since the MGC anterior border of the middle part of the cingulate gyrus is included in the HM's anterior part of the cingulate gyrus. Finally, the posterior border of MGC's middle part of the cingulate gyrus is included in the HM's posterior part of the cingulate gyrus. The superior border of the MGC middle cingulate gyrus and HM anterior cingulate gyrus is the same and the inferior is approximately the same.

The posterior and superior border differ. The HM anterior cingulate gyrus will have a larger volume than the MGC anterior cingulate gyrus.

3.1.4 Posterior cingulate gyrus

The anterior border defined by the MGC is the anterior limit of the splenium of the corpus callosum (plane in the anterior limit of the splenium of the corpus callosum). The HM defines the border as the anterior part of the cingulate gyrus, i.e. vertical lines drawn from the corpus callosum to sulcus cinguli at the midpoint of the greatest extension of the corpus callosum and the corpus callosum inferiorly. The borders will not coincide - the HM border will be further anterior. The region between the anterior border defined by the HM and the anterior border defined by the MGC will instead be covered by the middle part of the cingulate gyrus as defined by the MGC. The posterior border is partly defined by the sulcus subparietalis by both of the databases. Inferiorly the posterior border is a bit different since the HM uses the sulcus parieto-occipitalis as border, and the MGC uses the inferior projection from the inferior limit of the subparietal sulcus. This causes that the HM posterior border is located a bit

further posterior in the inferior part. Both of the MGC and the HM define the superior border as sulcus cinguli. The MGC also defines the subcallosal area border as callosal sulcus. The inferior border will be different between the two databases, since the MGC defines the subcallosal area border as the calcarine sulcus and the supracallosal area border as the callosal sulcus. The HM, however, defines this border as the most inferior slice on which the splenium corporis callosi is uninterrupted across its width. This will cause the inferior border in the HM atlases to be further superior than the subcallosal area border in MGC. The cingulate gyrus in the HM atlases will not extend to the end of the corpus callosum, which it will do in the MGC atlases. According to the figures in the HM protocol for region 50–83 obtained from a supplement to Gousias et al. (2008), the inferior border (along with the corpus callosum) is approximately the callosal sulcus in HM as well as in MGC, but the borders will be a bit tighter to the corpus callosum than in MGC since the corpus callosum borders are defined as where the cingulate gyrus is defined. The superior border of the middle cingulate gyrus defined by the MGC will coincide with the superior border of the HM posterior cingulate gyrus. The inferior borders are also approximately the same.

The anterior, posterior and inferior border differ, and the HM posterior cingulate gyrus region has a larger volume than the MGC posterior cingulate gyrus. The total volume of the cingulate gyrus defined by the MGC is smaller than the total volume of the HM region.

3.1.5 Fusiform gyrus

The MGC anterior border is a bit more consistent, since it is defined by the occipitotemporal sulcus, while the border in HM is defined as the first slice where the amygdala is seen. The anterior limit of the occipitotemporal sulcus and the amygdala is pretty close. The posterior border in the MGC is defined by a coronal plane in the anterior limit of the ventral bank of the parietooccipital sulcus, while in the HM the border is the most posterior slice where the hippocampus is seen. The posterior border for the occipital fusiform gyrus, however, is the posterior hemispheric margin. This means that the MGC occipital fusiform gyrus will extend much further posterior than the HM definition. The posterior border of the MGC fusiform gyrus is also further posterior than the HM border since the ventral bank of the parietooccipital sulcus is located further posterior then the last slice of the hippocampus. In addition, the fusiform gyrus will be divided into two parts by the MGC. The lateral and medial borders are reversed, i.e. the lateral border defined by the MGC is the occipitotemporal sulcus which is defined as the medial border in the HM protocol. The medial border defined by the MGC protocol is the collateral sulcus, which is defined as the lateral border in the HM protocol. This is just a mix up in the HM protocol. The MGC does not define the superior and inferior border.

The fusiform gyrus as defined by the MGC will overlap with the HM definition of the fusiform gyrus in some parts. They will not have the same exact borders, even though some are close to each other. The MGC fusiform gyrus will have a larger volume than the HM-defined fusiform gyrus.

3.2 Visualization of the overlap

The following section describes the results from the visualization of the overlaps in R. Only the atlases generated for the HM images were used.

3.2.1 Hippocampus

The visualization of the overlap between the HM and MGC's hippocampus region showed that the MGC region covered almost all of the HM region, see Figure 2. The MGC hippocampus extended

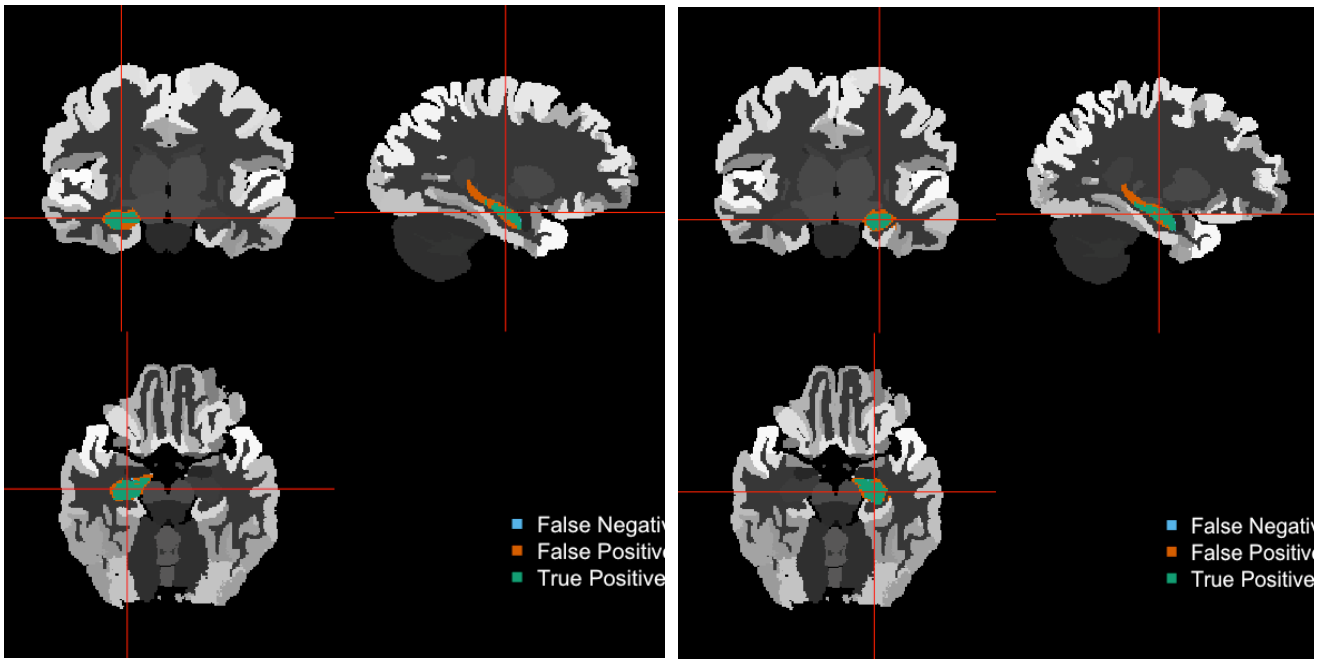

Figure 2: Visualization of the overlap between HM right and left hippocampus and the MGC-defined right and left hippocampus. Green (true positive) implies that the voxel belongs to both the HM and the MGC region. Blue (false negative) implies that the voxel belongs to the HM region but not the MGC region. Orange (false positive) implies that the voxel belongs to the MGC region but not the HM region.

further posterior and included the hippocampal tail while the HM excluded it. The HM hippocampus extended slightly further anterior. The HM hippocampus also extended further medial anterior. The HM hippocampus overlapped with the MGC amygdala anteriorly and inferiorly. It also overlapped with the inferior lateral ventricle anteriorly and at the superior border. The MGC cerebral white matter covered the HM hippocampus in almost every direction except for at the medial border. The overlap between the HM hippocampus and the MGC parahippocampal gyrus mostly occurred at the inferior and medial border. The HM parahippocampal and ambient gyrus overlapped with the MGC hippocampus by the inferior border both anteriorly and posteriorly. The MGC hippocampus overlapped with the lateral ventricle temporal horn superiorly and at the lateral border. The MGC hippocampus also overlapped with the HM amygdala in the anterior part.

3.2.2 Thalamus

The visualization of the overlap between the HM region thalamus and the MGC's region thalamus proper is illustrated in Figure 3 which shows that the regions overlap each other very well. The MGC thalamus proper extended a bit further in almost every direction, although in almost every image there were some voxels that belonged to the HM region that extended outside the MGC region. There was some inconsistency in the location of these voxels though. The MGC left thalamus proper especially extended further lateral and anterior wise. The HM thalamus overlapped the MGC ventral DC and the hippocampus at the inferior border. The MGC's cerebral white matter covered the HM thalamus close to all of the borders and seemed to overlap at the lateral, medial and superior border (but also the anterior and posterior). There was also some overlap with the MGC's third ventricle which seemed to be mostly at the medial border and mostly posterior and inferior. The MGC thalamus proper overlapped with the HM third ventricle by the medial border. The overlap with the insula posterior long gyrus occurred at the lateral border. The overlap between the MGC thalamus proper and the HM-defined lateral ventricle excluding temporal horn occurred at the superior border.

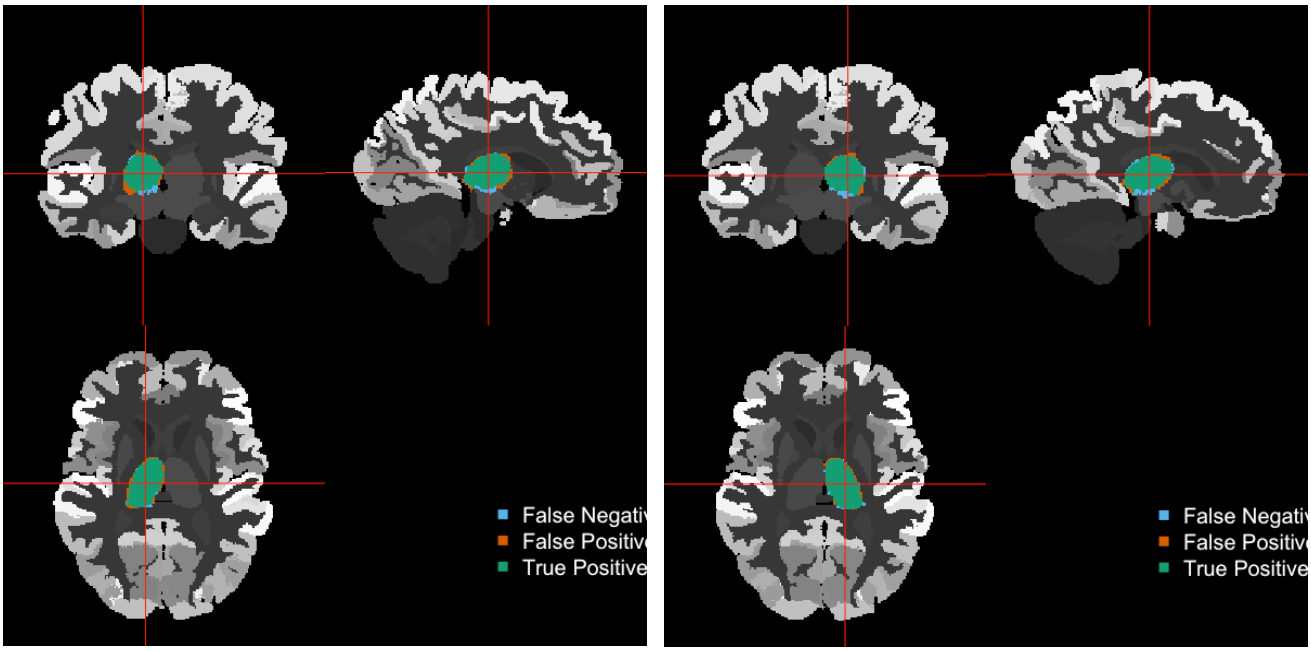

Figure 3: Visualization of the overlap between HM right and left thalamus and the MGC-defined right and left thalamus. Green (true positive) implies that the voxel belongs to both the HM and the MGC region. Blue (false negative) implies that the voxel belongs to the HM region but not the MGC region. Orange (false positive) implies that the voxel belongs to the MGC region but not the HM region.

3.2.3 Anterior cingulate gyrus

The anterior cingulate gyrus defined by the HM and the anterior cingulate gyrus defined by the MGC had some overlap, but the visualization of the overlap (see Figure 4) showed that the MGC region extended much further anterior than the HM region, while the HM region extended further posterior. The MGC region also extended further inferior in some parts, while the HM region extended a bit further lateral. The visualization of the overlap between the HM anterior cingulate gyrus and the MGC middle cingulate gyrus showed that the MGC region extended further posterior, while the HM region extended further lateral and anterior. The MGC superior frontal cortex medial segment overlapped the HM anterior cingulate gyrus at the superior (middle and anteriorly) and anterior border. Both the right and left HM anterior cingulate gyrus overlapped with MGC cerebral white matter at the lateral border. The MGC anterior cingulate gyrus overlap with the HM pre-subgenual frontal cortex occurred at the inferior border in the most anterior part, where there was no overlap with the HM anterior cingulate gyrus. The overlap with the HM pre-subgenual and subgenual frontal cortex occurred at the anterior and inferior part of the MGC anterior cingulate gyrus where the region was curved. The HM anterior cingulate gyrus did not include this curve. The overlap between the HM superior frontal gyrus and the MGC anterior cingulate gyrus occurred at the superior border along almost the whole MGC region and somewhat at the lateral border. But it also overlapped at the anterior inferior border where the region curved.

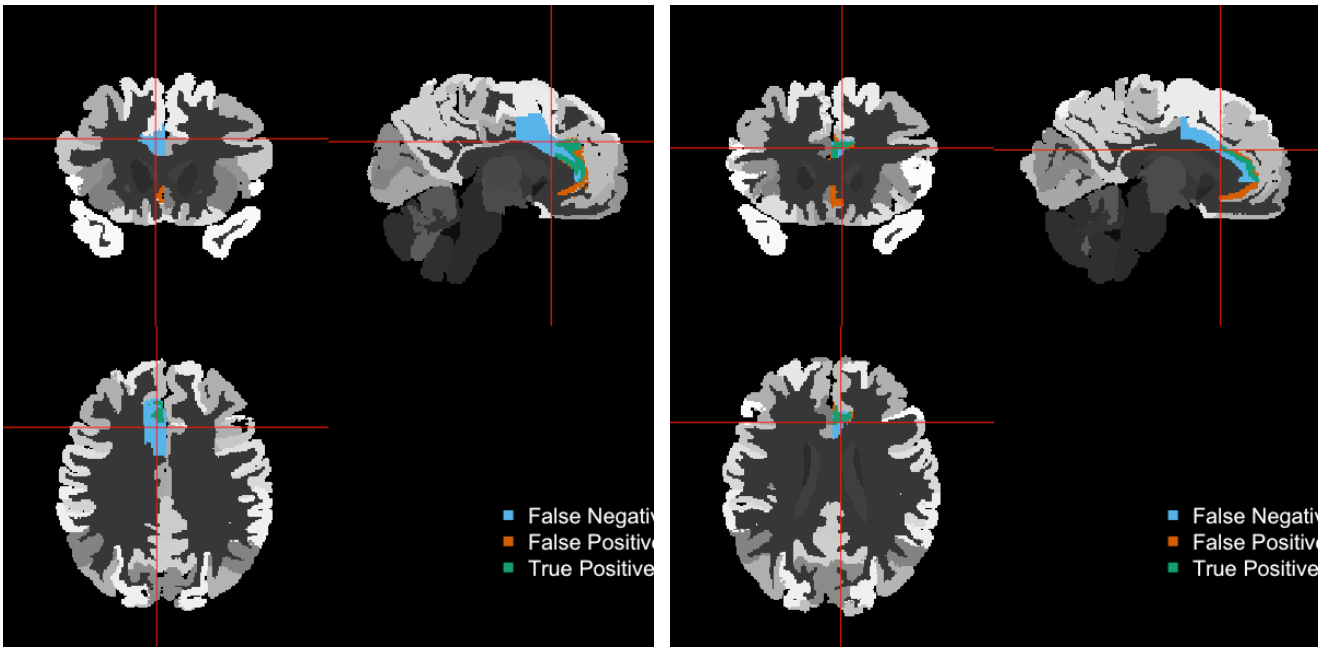

Figure 4: Visualization of the overlap between HM right and left anterior cingulate gyrus and the MGC-defined right and left anterior cingulate gyrus. Green (true positive) implies that the voxel belongs to both the HM and the MGC region. Blue (false negative) implies that the voxel belongs to the HM region but not the MGC region. Orange (false positive) implies that the voxel belongs to the MGC region but not the HM region.

3.2.4 Posterior cingulate gyrus

The MGC posterior cingulate gyrus overlapped with the HM posterior cingulate gyrus. The HM region extended further anterior while the MGC region extended further posterior and inferior. The HM region also had parts in the centre of the region that were not overlapped by the MGC posterior cingulate gyrus region (see Figure 5). Most of this part was instead overlapped by the MGC cerebral white matter. The MGC cerebral white matter also covered the HM posterior cingulate gyrus at the lateral and inferior border which was where the overlap occurred. The overlap also occurred in the centre of the HM region and not just by the borders and covered a large part of the region. The HM posterior cingulate gyrus overlapped with the MGC middle cingulate gyrus in the anterior part. The MGC posterior cingulate gyrus overlapped with the HM superior temporal lobe in the inferior and posterior part. The overlap between the MGC posterior cingulate gyrus and the HM superior parietal gyrus occurred at the superior border but also inferiorly posterior, superior to the overlap with the superior temporal lobe. It also overlapped at the lateral border somewhat. These overlaps occurred inferior to the location of the HM posterior cingulate gyrus inferior posterior border.

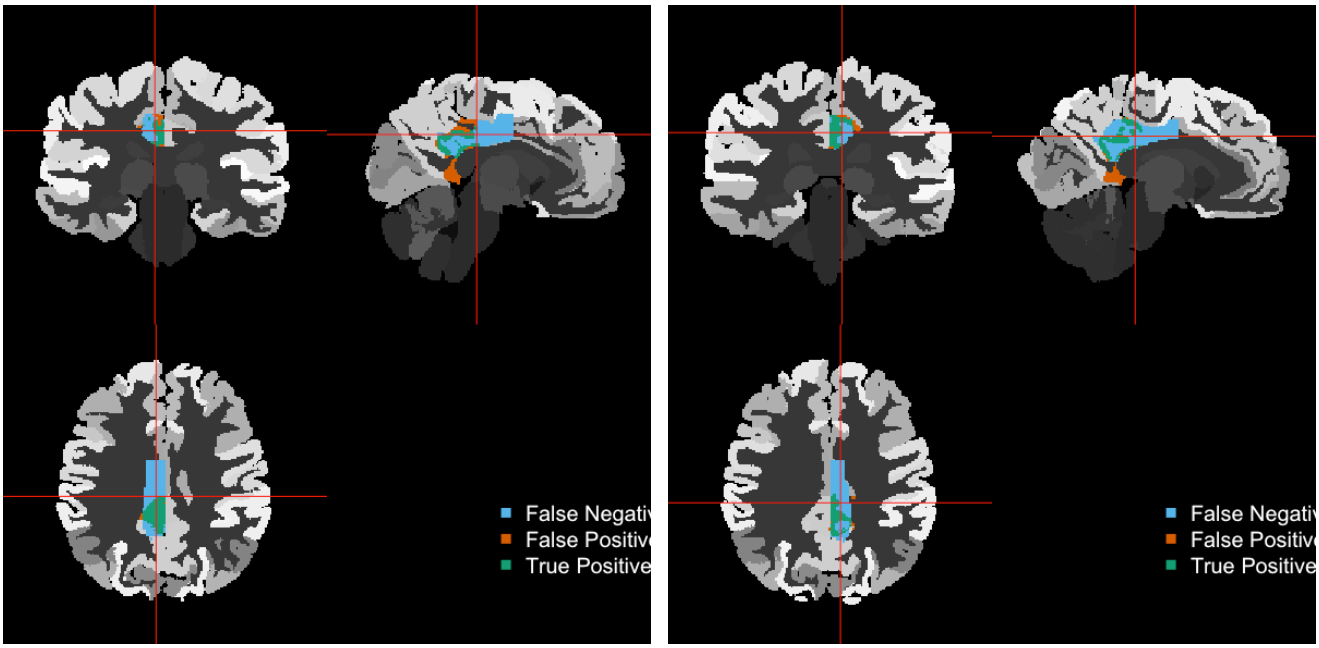

Figure 5: Visualization of the overlap between HM right and left posterior cingulate gyrus and the MGC-defined right and left posterior cingulate gyrus. Green (true positive) implies that the voxel belongs to both the HM and the MGC region. Blue (false negative) implies that the voxel belongs to the HM region but not the MGC region. Orange (false positive) implies that the voxel belongs to the MGC region but not the HM region.

3.2.5 Fusiform gyrus

The visualization of the overlap between the HM fusiform gyrus and MGC fusiform gyrus is illustrated in Figure 6 which showed that the HM region extended further anterior while the MGC region extended further posterior. The HM region also extended further inferior in the anterior part. The overlap between the HM fusiform gyrus and the MGC inferior temporal gyrus occurred at the anterior border and at the lateral and more towards the inferior border than the superior. The HM fusiform gyrus and the MGC parahippocampal gyrus overlapped at the medial border. The MGC fusiform gyrus had a large overlap with the HM posterior temporal lobe in the posterior part. This overlap covered almost half of the MGC region in some slices and the posterior part that was not overlapped by the HM fusiform gyrus. However, it did not cover the most posterior part of the MGC region. This part was instead overlapped by the HM lateral remainder occipital lobe. The MGC fusiform gyrus also had some overlap with the HM parahippocampal and ambient gyrus which occurred at the superior border anteriorly and at the medial border. The overlap between the MGC fusiform gyrus and the HM middle and inferior temporal gyrus occurred at the lateral border and somewhat anterior.

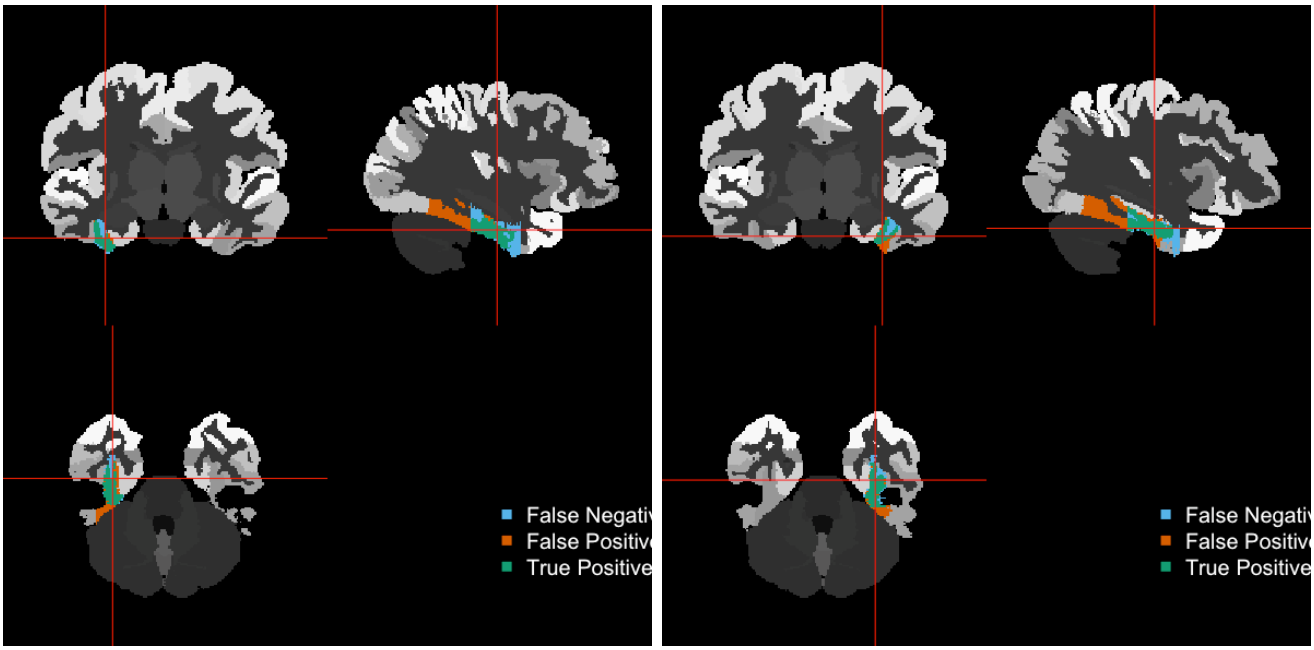

Figure 6: Visualization of the overlap between HM right and left fusiform gyrus and the MGC-defined right and left fusiform gyrus. Green (true positive) implies that the voxel belongs to both the HM and the MGC region. Blue (false negative) implies that the voxel belongs to the HM region but not the MGC region. Orange (false positive) implies that the voxel belongs to the MGC region but not the HM region.

3.3 Global comparison of the region names

73 of the HM regions had matching regions names with the MGC regions and 86 of the MGC regions had matching region names with the HM regions (see Appendix Table 5 and Table 6). The main differences between the defined regions from the databases were the subdivisions of regions and inclusion of different gyri. MGC generally defined smaller and more specific regions, like the entorhinal area and the supplementary motor cortex, that can be linked to specific brain functions. The MGC also generally had smaller subdivisions: the inferior frontal gyrus, for example, was divided into three parts instead of one, and the occipital lobe which was divided into the occipital gyrus, occipital pole and cuneus instead of cuneus and lateral remainder occipital lobe. The insula, on the other hand was divided into six parts in the HM protocol but only into two parts in the MGC protocol. The temporal lobe was also further subdivided in the HM, and the regions had names that explained which part of the lobe it covered, while in the MGC smaller parts like the temporal pole and planum temporale were defined instead. Some large regions that were defined in the MGC but did not have a corresponding region in the HM were the ventral DC, vessel (vessels inferior the putamen), basal forebrain, cerebral exterior and white matter, CSF, the 4th ventricle, and the central/frontal/parietal operculum. Some regions that only the HM protocol defined were the corpus callosum, substantia nigra, subgenual frontal cortex, and pre-subgenual frontal cortex. The extra regions in MGC were mostly cortical and a few subcortical regions.

3.4 Most frequent coinciding combinations

40 and 43 of the HM regions had the matching MGC region as the most frequent coinciding region in the HM images and the MGC images respectively, and 76 of the MGC regions had the matching HM region as the most frequent coinciding region in both the HM and MGC images. Many of the regions

defined by the HM protocol that did not have a matching region in the MGC protocol had cerebral white matter as the most frequent coinciding MGC region. This means that cortical regions only include the actual cortical grey matter in the MGC, whereas HM includes adjacent white matter as part of the region. The MGC-defined ventral DC, basal forebrain and 4th ventricle had the HM background as the most frequent coinciding. Both atlas databases leave certain brain regions unclassified (assigned the background label) and do not cover the brain in its entirety.

Some HM regions that had a ratio between the number of voxels that belonged to the combination with the second and first most frequent coinciding region that was close to one (> 0.75) were the precentral gyrus, insula anterior long gyrus, insula anterior short gyrus and middle frontal gyrus. The ratio between the number of voxels that belonged to combination with the third and first most frequent coinciding region was small (< 0.20) for these regions. These regions also had cerebral white matter or the expected MGC region as the first and second most frequent region. There were also some HM regions that had a very small ratio between the number of voxels that belonged to the combination with the second and first most frequent region. These were, for example, the thalamus, the hippocampus, the lateral ventricle excluding temporal horn, and the brainstem excluding substantia nigra. The analysis of the ratios showed for example that the HM´s corpus callosum roughly matched half the right cerebral white matter and half the left cerebral white matter as defined by the MGC.

The MGC also had some regions where the ratio between the number of voxels that belonged to the combination with the second and first most frequent coinciding region was close to one and the ratio between the combinations with the third and first most frequent region was small. These regions were for example, the right calcarine cortex, the middle cingulate gyrus, cerebellar vermal lobules I–V, cerebellar vermal lobules VI–VII and cerebellar vermal lobules VIII–X. The middle cingulate gyrus roughly matched half HM posterior and half HM anterior cingulate gyrus. Regions which had a very small ratio between the number of voxels that belonged to the combination with the second and first most frequent region were especially regions which had an obvious corresponding HM region, for example the lateral ventricle, the middle frontal gyrus and the precentral gyrus. The MGC had more regions with smaller ratios than the HM, while the HM had more regions with high ratios.

The ratio between the number of voxels that belonged to the combination with the second and first most frequent coinciding region was generally higher for those regions where there was an inconsistency regarding which combination of regions coincided most frequently. The mean was generally higher for the HM regions that had a matching region in the other protocol, but did not have this region as the most frequent or regions that were divided into more regions by the other protocol, for example, the HM's anterior and posterior cingulate gyrus, the middle and inferior temporal gyrus and the subcallosal area. There was a correlation between the ratios between the region pairs.

Some regions that had an inconsistency between the images still had quite a low mean ratio between the number of voxels that belonged to the combination with the second and first most frequent coinciding regions. This usually occurred when the investigated region was a subdivision of the region that was the most frequent coinciding. For example, MGC's postcentral gyrus medial segment and the orbital part of the inferior frontal gyrus that had the HM postcentral gyrus and inferior frontal gyrus as the most frequent coinciding.

3.5 Jaccard coefficient

The calculation of the Jaccard coefficient showed that the greatest overlap occurred between the regions that had matching names with regions in the other protocol. The standard deviation of the Jaccard coefficient was approximately of the same magnitude for almost all of the 120 largest overlaps. However there were some overlaps that were noticeable, for example the overlap between the HM lateral ventricle temporal horn and the MGC inferior lateral ventricle (only the right region in the HM images but both in the MGC images) as well as the overlap between the HM lateral orbital gyrus and the MGC lateral orbital gyrus (only the left region in the HM images but both in the MGC images), which both had a high standard deviation compared to the mean. The largest overlap occurred between the HM lateral ventricle excluding temporal horn and the MGC lateral ventricle and the second largest occurred between the putamen regions. These overlaps had a small standard deviation. Approximately 120 regions had an overlap larger than 10 % and most of these had a small standard deviation compared to the mean and were regions that overlapped with their matching region. All of these overlaps occurred in every combined atlas. Regions that had an overlap larger than 10 % and a small variation, but were not predicted based on their names were for example the overlap between the HM superior parietal gyrus and the MGC precuneus, overlap between the HM posterior temporal lobe and the MGC middle temporal gyrus, and the overlap between the HM cuneus and MGC calcarine cortex. There were multiple overlaps that just occurred in one of the combined atlases of the HM images, for example overlap between the right HM-defined cuneus and right MGC-defined posterior cingulate gyrus, between the left HM lingual gyrus and the left MGC inferior occipital gyrus and between the left HM subcallosal area and the left MGC ventral DC. The smaller overlaps had larger deviations compared to the mean. There was a good correlation between the region pairs and the mean Jaccard coefficient.

3.6 Volumes

The HM regions generally had larger volumes compared to the corresponding MGC regions, although there were exceptions where MGC regions were almost twice the size of the corresponding HM regions. Out of the 86 MGC regions that had matching HM regions, the ratio between volumes was greater than one for 74 both when using the HM and the MGC images. 51 and 53 out of the 73 HM regions that had a matching MGC region had a ratio that was greater than one when using the HM images and the MGC images respectively. The ratio between the HM volume of the background and the MGC background was very close to one. The ratios were consistent between the right and left hemisphere and were of the same order of magnitude. Some regions that had very high ratios even though they had matching region names and were not subdivisions were the anterior orbital gyrus, the straight gyrus, the superior frontal gyrus and the precentral gyrus. The subcallosal area, the fusiform gyrus, and the hippocampus are examples of regions that had small ratios. This was determined from the ratios calculated when all of the atlases were added together.

Most of the region volumes had quite large variation. 18 regions in the right side of the brain in the HM images and 24 in the MGC images had a larger volume than their counterpart out of 46 region pairs. There was no significant difference between the CoV between small and large regions. Some HM regions that had a high CoV were the subcallosal area, the pre-subgenual frontal cortex, and the lateral ventricle excluding temporal horn. Regions that had small CoV were the HM background, cerebellum, and putamen in the HM images and the HM background, cerebellum, and parahippocampal and ambient gyrus in the MGC images. The largest HM regions were the

cerebellum, superior frontal gyrus, and middle frontal gyrus. The smallest regions were the nucleus accumbens, subcallosal area, and substantia nigra.

33 and 32 right side regions had larger volumes out of 64 MGC region pairs in the HM images and the MGC images, respectively. The CoV did not correlate with the volume of the region, except for very large or very small regions. The very small regions had a large CoV and the very large volumes had a small CoV. Some MGC regions that had a high CoV were the lateral ventricle, vessel and third ventricle. The optic chiasm had a CoV that was almost one and had a very small volume of about 30 voxels. Regions that were among the smallest in terms of CoV were brainstem and the cerebellum exterior. The largest MGC regions were the cerebral white matter, the cerebellum exterior and the middle frontal gyrus. The smallest MGC regions were the vessel, optic chiasm, accumbens area, and the inferior lateral ventricle.

The CoV of the volume ratio between the HM regions and the MGC regions did not depend on the volume of the regions. There was some correlation between the CoV of the region volume and the CoV of the volume ratio. If the CoV for the region volume was large for one of the regions, it tended to be large for the ratio as well with some exceptions. There was a clear correlation between the volume ratios for each region pair.

3.7 Quantitative comparison of the four chosen regions

The results from the calculation of the Jaccard coefficient, volume ratio, number of overlaps and the most frequent coinciding regions for the four regions predictions were made for are found in Table 1, Table 2, Table 3 and Table 4. The first row for each region is the result calculated from the HM images, the second row is the result from the MGC images and if there only is one row the result was the same for all of the images. Table 1 and Table 2 show the first, second and third most frequent coinciding regions for the four investigated regions. These were calculated from when all of the atlases were added together. The mean volume ratio and the mean Jaccard coefficient between the corresponding HM and MGC regions can be found in Table 3. The two columns on the far right of the table represent the total number of different regions from the other protocol the investigated region overlapped with. The largest overlaps according to the calculation of the mean Jaccard coefficient for the four regions are shown in Table 4.

Table 1: First, second and third most frequent MGC coinciding regions and the corresponding ratios for the HM regions predictions were made for, calculated when all of the images were added. The first row for each region is results obtained from the HM images and the second row is results obtained from the MGC images.

HM Regions (first row = HM images, second row = MGC images)	**First**	**Second**	**Third**	**Ratio (second / first)**	**Ratio (third / first)**
Right Hippocampus	Hippocampus	Amygdala	Inferior lateral ventricle	0.053	0.014
				0.036	0.030

Left Hippocampus	Hippocampus	Amygdala	Cerebral white matter	0.046	0.018
		Inferior lateral ventricle		0.019	0.017
Right Thalamus	Thalamus Proper	Ventral DC	Cerebral white matter	0.045	0.044
		Cerebral white matter	Ventral DC	0.101	0.050
Left Thalamus	Thalamus Proper	Ventral DC	Cerebral white matter	0.056	0.028
		Cerebral white matter	Ventral DC	0.076	0.060
Right Anterior cingulate gyrus	Anterior cingulate gyrus	Cerebral white matter	Middle cingulate gyrus	0.822	0.768
		Middle cingulate gyrus	Cerebral white matter	0.753	0.718
Left Anterior cingulate gyrus	Anterior cingulate gyrus	Cerebral white matter	Middle cingulate gyrus	0.871	0.697
		Middle cingulate gyrus	Cerebral white matter	0.707	0.705
Right Posterior cingulate gyrus	Posterior cingulate gyrus	Cerebral white matter	Middle cingulate gyrus	0.941	0.837
		Middle cingulate gyrus	Cerebral white matter	0.944	0.899
Left Posterior cingulate gyrus	Posterior cingulate gyrus	Cerebral white matter	Middle cingulate gyrus	0.861	0.669
				0.782	0.775
Right Fusiform gyrus	Fusiform gyrus	Cerebral white matter	Inferior temporal gyrus	0.360	0.144

				0.357	0.157
Left Fusiform gyrus	Fusiform gyrus	Cerebral white matter	Inferior temporal gyrus	0.418	0.300
				0.397	0.318

Table 2: First, second and third most frequent coinciding HM regions and the corresponding ratios for the MGC regions predictions were made for, calculated when all of the images were added. The first row for each region is results obtained from the HM images and the second row is results obtained from the MGC images.

MGC Regions (first row = HM images, second row = MGC images)	**First**	**Second**	**Third**	**Ratio (second / first)**	**Ratio (third / first)**
Right Hippocampus	Hippocampus	Posterior temporal lobe	Parahippocampal and ambient gyrus	0.340	0.200
				0.405	0.279
Left Hippocampus	Hippocampus	Posterior temporal lobe	Parahippocampal and ambient gyrus	0.340	0.218
				0.423	0.270
Right Thalamus proper	Thalamus	Background	Brainstem excluding substantia nigra	0.225	0.010
			Lateral ventricle excluding temporal horn	0.128	0.005
Left Thalamus proper	Thalamus	Background	Brainstem excluding substantia nigra	0.240	0.006
			Lateral ventricle excluding temporal horn	0.125	0.005
Right Anterior cingulate gyrus	Anterior cingulate gyrus	Pre-subgenual frontal cortex	Subgenual frontal cortex	0.192	0.164

		Superior frontal gyrus	Pre-subgenual frontal cortex	0.167	0.147
Left Anterior cingulate gyrus	Anterior cingulate gyrus	Pre-subgenual frontal cortex	Superior frontal gyrus	0.254	0.243
		Superior frontal gyrus	Pre-subgenual frontal cortex	0.269	0.225
Right Posterior cingulate gyrus	Posterior cingulate gyrus	Posterior temporal lobe	Superior parietal gyrus	0.331	0.259
				0.333	0.266
Left Posterior cingulate gyrus	Posterior cingulate gyrus	Posterior temporal lobe	Superior parietal gyrus	0.281	0.187
				0.317	0.225
Right Fusiform gyrus	Posterior temporal lobe	Fusiform gyrus	Lateral remainder occipital lobe	0.659	0.128
			Background	0.594	0.048
Left Fusiform gyrus	Posterior temporal lobe	Fusiform gyrus	Lateral remainder occipital lobe	0.619	0.166
				0.563	0.072

Table 3: Mean volume ratio, mean Jaccard coefficient and number of different regions the investigated region overlapped with. The first row for each region is results obtained from the HM images and the second row is results obtained from the MGC images.

Regions (first row = HM images, second row = MGC images)	**Volume ratio mean ± SD (CoV)**	**Jaccard Coefficient mean ± SD (CoV)**	**Number of MGC regions the HM region overlapped with**	**Number of HM regions the MGC region overlapped with**
Right Hippocampus	0.583 ± 0.051 (0.088)	0.507 ± 0.034 (0.067)	10	16
	0.553 ± 0.040 (0.073)	0.480 ± 0.033 (0.068)	11	13
Left Hippocampus	0.562 ± 0.048 (0.086)	0.496 ± 0.037 (0.075)	9	14
	0.524 ± 0.041 (0.079)	0.475 ± 0.034 (0.073)	10	13
Right Thalamus	0.877 ± 0.079 (0.090)	0.740 ± 0.028 (0.038)	14	14
	1.01 ± 0.066 (0.065)	0.764 ± 0.023 (0.031)	15	10
Left Thalamus	0.867 ± 0.075 (0.086)	0.735 ± 0.032 (0.044)	14	18
	1.01 ± 0.053 (0.053)	0.774 ± 0.019 (0.025)	15	11
Right Anterior cingulate gyrus	1.94 ± 0.429 (0.222)	0.274 ± 0.048 (0.176)	12	19
	1.96 ± 0.437 (0.223)	0.298 ± 0.045 (0.152)	14	16
Left Anterior cingulate gyrus	1.64 ± 0.391 (0.239)	0.255 ± 0.047 (0.182)	11	19
	1.67 ± 0.457 (0.273)	0.285 ± 0.047 (0.163)	15	16
Right Posterior cingulate gyrus	1.64 ± 0.265 (0.162)	0.255 ± 0.032 (0.124)	14	19

	1.82 ± 0.441 (0.243)	0.255 ± 0.045 (0.178)	16	14
Left Posterior cingulate gyrus	1.60 ± 0.173 (0.108)	0.295 ± 0.037 (0.126)	13	18
	1.70 ± 0.346 (0.204)	0.283 ± 0.043 (0.151)	17	14
Right Fusiform gyrus	0.540 ± 0.088 (0.162)	0.273 ± 0.048 (0.175)	12	11
	0.546 ± 0.082 (0.150)	0.276 ± 0.052 (0.189)	11	13
Left Fusiform gyrus	0.579 ± 0.142 (0.245)	0.248 ± 0.055 (0.224)	12	10
	0.573 ± 0.107 (0.186)	0.246 ± 0.072 (0.292)	12	13

Table 4: The regions which the investigated regions had the largest overlap with according to the mean Jaccard coefficient. The first row for each region is results obtained from the HM images and the second row is results obtained from the MGC images. If there only is one row the result was the same for all images. The oblique expresses the regions corresponding to the right / left region if they were different.

Regions	**First**	**Second**	**Third**
HM Right & Left Hippocampus	Hippocampus	Amygdala	Inferior lateral ventricle
		Inferior lateral ventricle	Amygdala
MGC Right & Left Hippocampus	Hippocampus	Lateral ventricle temporal horn	Parahippocampal and ambient gyrus
		Parahippocampal and ambient gyrus	Lateral ventricle temporal horn
HM Right & Left Thalamus	Thalamus proper	Ventral DC	Hippocampus
			Cerebral white matter / Hippocampus
MGC Right & Left Thalamus proper	Thalamus	Third ventricle	Insula posterior long gyrus
			Lateral ventricle excluding temporal horn

HM Right & Left Anterior cingulate gyrus	Anterior cingulate gyrus	Middle cingulate gyrus	Supplementary motor cortex / Superior frontal gyrus medial segment
MGC Right & Left Anterior cingulate gyrus	Anterior cingulate gyrus	Pre-subgenual frontal cortex	Subgenual frontal cortex
MGC Right & Left Middle cingulate gyrus	Posterior cingulate gyrus	Anterior cingulate gyrus	Left anterior cingulate gyrus / Right posterior cingulate gyrus
			Corpus Callosum / Superior frontal gyrus
HM Right & Left Posterior cingulate gyrus	Posterior cingulate gyrus	Middle cingulate gyrus	Precuneus
MGC Right & Left Posterior cingulate gyrus	Posterior cingulate gyrus	Posterior temporal lobe	Superior parietal gyrus
HM Right & Left Fusiform gyrus	Fusiform gyrus	Inferior temporal gyrus	Parahippocampal gyrus
MGC Right & Left Fusiform gyrus	Fusiform gyrus	Posterior temporal lobe	Parahippocampal and ambient gyrus / Lateral remainder occipital lobe
			Anterior temporal lobe medial part

The hippocampus can be used as an instructive example of how to read the tables. According to Table 1, the HM-defined right hippocampus had the right hippocampus, amygdala, and inferior lateral ventricle defined by the MGC as the first, second and third most frequent coinciding region whether the HM and MGC images were used. The left hippocampus defined by the HM had the left hippocampus, amygdala and cerebral white matter as the first, second, and third most frequent coinciding MGC region when the HM images were used. The inferior lateral ventricle defined by the MGC was the second most frequent coinciding when the MGC images were used. Table 2 shows that the MGC label defining the hippocampus had the hippocampus, posterior temporal lobe, and parahippocampal and ambient gyrus defined by the HM as the first, second, and third most frequent

regions for both the left and right region regardless of which images were used. The 30 right HM hippocampi that were analysed overlapped with 10 and 11 different MGC regions in total using the HM images and the MGC images, respectively (see Table 3). The left HM hippocampus overlapped with 9 and 10 different MGC regions while the right MGC-defined hippocampus overlapped 16 and 13 different HM regions and the left overlapped with 14 and 13 when the HM images and the MGC images were used, respectively. Table 3 also shows that the mean Jaccard coefficient for the right hippocampus was 0.507 ± 0.034 and 0.496 ± 0.037 for the left hippocampus using the HM images. The mean volume ratio was 0.553 ± 0.040 for the right hippocampus and 0.524 ± 0.041 for the left, calculated using the MGC images, which means that the MGC-defined hippocampus was almost twice the size as the HM-defined hippocampus. Table 4 shows that the HM-defined hippocampus had the largest overlaps with the MGC hippocampus, amygdala and inferior lateral ventricle and that the MGC-defined hippocampus had the largest overlap with the HM hippocampus, lateral ventricle temporal horn and parahippocampal and ambient gyrus according to the mean Jaccard coefficient calculated from the HM images. When the MGC images were used the second and third largest overlap were reversed compared to the result from the HM images.

3.8 Evaluation of the quantitative comparison

The quantitative comparison of the hippocampus confirmed the most important predictions. For example, it was confirmed that the MGC hippocampus had a larger volume than the HM-defined hippocampus, that both regions overlapped with each other and that each region had some overlap with the nearby regions where the border definitions differed. The quantitative comparison also revealed additionally overlaps that were not predicted, for example that the MGC hippocampus overlapped with the posterior temporal lobe defined by the HM. But on examining the protocol for the posterior temporal lobe one finds that it is located close to the hippocampus and it is reasonable that the regions overlap.

The quantitative comparison confirmed that the thalamus regions overlapped each other very well, based on both the Jaccard coefficient and the volume ratio. It also confirmed the predicted overlaps and revealed some unpredicted, for example the overlap between the MGC thalamus proper and the HM-defined brainstem excluding substantia nigra. The volume ratio also revealed that the MGC-defined thalamus proper was larger than the HM-defined thalamus region when using the HM images but not when using the MGC images.

The results from the analysis of the most frequent coinciding regions for the HM anterior cingulate gyrus confirmed that the largest overlaps occurred between the MGC anterior cingulate gyrus and middle cingulate gyrus which was predicted. It also revealed for example, that the HM anterior cingulate gyrus overlapped with the supplementary motor cortex defined by the MGC which was not predicted. The predicted volume difference between the corresponding regions was confirmed by the volume ratio.

The quantitative comparison of the posterior cingulate gyrus confirmed the predictions, both the overlaps and the volume difference. It also revealed overlaps that was not predicted, for example the overlap between the MGC posterior cingulate gyrus and the HM postcentral gyrus. The calculation of the ratio between the number of voxels that belonged to the combination with the second and first most frequent coinciding region for the HM posterior cingulate gyrus was high which meant that the second and first most frequent coinciding MGC region covered almost the same amount of the HM

region. There was also an inconsistency regarding which MGC region coincided most frequently with the HM posterior cingulate gyrus which could not be predicted. The volume ratio between the total volume of the cingulate gyrus was larger than one which confirmed the prediction that the HM region is larger than the MGC volume.

The predictions for the fusiform gyrus was confirmed by the quantitative comparison, both the predicted overlaps and volume ratio. The overlap between the HM-defined fusiform gyrus and the MGC cerebral white matter was not predicted but could be explained by the lack of MGC descriptions of the inferior and superior border. The quantitative comparison also revealed that the MGC-defined occipital fusiform gyrus did not have any overlap with the HM fusiform gyrus and that the most frequent coinciding region with the MGC fusiform gyrus was not the expected HM fusiform gyrus but the HM posterior temporal lobe.

To summarize, the quantitative comparison confirmed most of the predictions and revealed multiple additional overlaps and insights that could not be predicted just based on studying the protocols. The result from the HM images and the MGC images were very similar and revealed essentially the same additional insights.

4 Discussion

The strength and novelty with the proposed method in this project are that it is composed of both a qualitative and a quantitative comparison between two atlas databases. It quantifies the differences in the delineation protocols and the differences and similarities in locations between regions. Another strength is that it could be applied to two different atlases, that have different number of regions, and be applied on both manually and automatically generated atlases. The quantitative method could be enough to enable mapping between two atlases and give details about the relationship between regions that overlap.

A weakness with the method is that the matrix dimensions of the overlapped atlases need to coincide. Another weakness is that the quantitative method is lacking a way to describe the location of the overlaps. The visualization of the overlaps, described in the qualitative comparison, is thus an essential part of the comparison, given that the location of the overlap is sought after.

With both the qualitative and quantitative comparisons it is possible to determine and find the difference between each region separately. This enables mapping between two brain atlases, and suggest a solution to the atlas concordance problem. The nomenclature problem is partly solved by the global comparison, which could be extended further by studying the protocols more thoroughly to understand if two regions with different names actually define the same region. The quantitative comparison and the visualization made it possible to describe differences and discrepancies between regions in detail.

4.1 Qualitative comparison

The four regions that the predictions were made for were chosen because they have important brain functions and had a corresponding region with matching name in each protocol. There were several factors that made predictions from the protocols difficult. One was that the HM and MGC used different nomenclature for the regions. The nomenclature problem was partly solved by the global comparison where region names were matched. Still, it was quite difficult to match some names due to differences in nomenclature. Some regions were subdivisions of a bigger part in the other atlas, and therefore the names were nearly identical, apart from the part of the name that explained the subdivision. This made it difficult to know how much of the region the subdivision covered, but this problem could perhaps be solved by studying the protocols. Some regions may also have had obvious overlap even though they did not have similar names, for example the temporal pole defined by the MGC and the anterior temporal lobe medial part defined by the HM. This could also be solved by studying the protocols more closely.

Another issue that affected the predictions was that some of the descriptions of the borders in the protocols lacked specificity and were open to interpretation. This made it difficult to compare the borders and draw conclusions about how the borders were related. The regions were also delineated in different slice orientations, which could affect how the borders were defined and described. For some regions the borders in some directions were not defined at all, which made it almost impossible to compare the affected region borders. For example, the MGC did not define the superior and inferior border for the fusiform gyrus. The figures that showed some of the regions in the protocols that were used to clarify some borders were very difficult to compare between the two databases, partly because they illustrated the regions in different orientations, but also because the MGC only described a

schematic of the labelling of the cortical regions. Some of the figures that should assist delineation of the subcortical regions were also very difficult to interpret because the adjacent regions were not named, and everything was drawn with the same colour. This part of the study was time-intensive, even though it was restricted to four regions to suit the scope of the project.

Another factor that could have affected the analysis is that the MGC document an internal procedure, whereas the HM invite collaboration and further protocol development. This could mean that the HM protocol may have changed for some regions during the development of the final protocol, potentially difficult to detect. The collaboration strategy of the HM protocol also causes it to describe the borders for different regions differently and also delineate the regions differently. For example, different collaborations use different delineation tools that could affect the precision of the borders and different amount of white or grey matter could be included in the regions depending on who did the delineation. The four regions chosen for analysis of the quantitative and qualitative method developed in this study should not be much affected by this as they were all described in one protocol produced at the same time.

The visualization of the four regions and their overlaps done in R using the *ortho_diff* function illustrated only one slice of the image, which made it difficult to see the location of the overlap between some region if the overlap did not occur in that slice. But for most of the overlaps it was possible to at least localise at which border the overlap should occur. Another problem was that the overlap with the background could not be visualised because if the background was plotted together with the region, all of the region was marked as to have an overlap with the background which was not true for any of the regions. The visualization could be done with the MGC images as well to validate the results from the HM images since the overlap should be similar independent of image, but it did not fit the timeframe.

4.2 Most frequent coinciding combinations

The calculation of the most frequent coinciding region can be used to get an overview of which regions that definitely have a corresponding region in the other protocol. The ratio between the number of voxels that belonged to the combination with the second and first most frequent coinciding region could be used to get an understanding of the relationship between the areas of the investigated region that was covered by the first, second and third most frequent coinciding regions. With a good knowledge of the brain anatomy it could be sufficient with the information given by these parameters to know approximately where the overlap occurs based on the region names.

The HM had a smaller number of regions that had the expected region as the most frequent coinciding region compared to the MGC. The reason behind this could be that the MGC had more and smaller subdivisions that had a matching region that was obvious and much larger. This could also be explained by the fact that the ratio between the volumes was greater than one for the clear majority of the MGC regions.

If the ratio between the number of voxels that belonged to the combination with the second and first most frequent coinciding region was close to one, it meant that these two regions roughly covered the same amount of the investigated region. If additionally, the ratio between the third and first most frequent coinciding region was small it suggested that the first and second most frequent region roughly covered the region by 50 % each. If the ratio between the second and first most frequent coinciding regions instead was small, it implied that the most frequent coinciding region roughly

covered the clear majority of the investigated region. An interesting observation from this calculation was that the corpus callosum roughly matched half right cerebral white matter and half left cerebral white matter which could be expected since the MGC did not define a region that matched the HM's corpus callosum. The corpus callosum is an unpaired symmetric structure that crosses the midline, which could explain why it overlapped with both right and left cerebral white matter. The reason why regions that had no matching region in the other protocol often had ratios that were high was probably that these regions overlapped with multiple regions. For these regions there was also some inconsistency between each of the 30 combined atlases, because there was no region that overlapped the clear majority of the investigated region.

The reason why some regions that had a region that was a subdivision of the investigated region as the most frequent still had a small ratio between the second and first most frequent region could be that the subdivision covers a part of the region that is large, but not as large as a region that had their matching region as the most frequent coinciding. There also was some inconsistency regarding which combination of regions coincided most frequently between the combined atlases for these regions, which probably was due to multiple subdivisions covering parts of the region. That there was a correlation between a high ratio between the second and first most frequent coinciding region and an inconsistency between the combined atlases regarding which region was the most frequent was expected. This, because if the ratio was high, the first and second most frequent region covered almost the same amount of area of the investigated region which could cause that in some combined atlases the second most frequent region was the first most frequent coinciding.

4.3 Jaccard Coefficient

The mean Jaccard coefficient and the standard deviation can be used to understand which overlaps occur, how large they are and how much they differ between the combined atlases. The standard deviation is a measurement that indicate if the regions overlap the same amount in every image. An advantage with the Jaccard coefficient is that it is a well-established measurement to denote overlap and is used in other studies that have quantified differences between for example, manually segmented and automatically segmented atlases. This makes it possible to compare results between studies. The Jaccard coefficient was also used because it is easily available and easily calculated.

The standard deviation of the Jaccard coefficient for the 120 largest overlaps was small compared to the mean, which implies that these overlaps were about the same size in every combined atlas they occurred in which could imply that these were not just random occurrences. Especially since all of them occurred in every combined atlas. There were combinations that did have coefficients greater than 10 % and a small variation but was not expected since they did not have matching region names, which could be evidence of systematic mismatches because the overlap was quite large and occurred in every combined atlas. The overlap between the HM's superior parietal gyrus and the MGC's precuneus seemed to be a systematic occurrence because of its small variation but it was not predicted in the global comparison between the regions because of their completely different names. The precuneus is, however, a portion of the superior parietal lobule, which explained the overlap. This phenomenon could though also appear when the overlaps only occurred in a few combined atlases because the variation between a few numbers of combined atlases will most likely be smaller than the variation between all of the combined atlases. This was not the case in this overlap, though, because all of the overlaps that had a Jaccard coefficient larger than 10 % occurred in all of the combined atlases. Overlaps that only occurred in one or a few images were probably just random mismatches

and could be due to the automatic segmentation. The largest overlaps all seemed to be systematic, which was expected since almost all of them overlapped with their matching region from the other protocol or their bordering regions. Overall though it seemed like unexpected overlaps were small and varied more in proportion to the mean between the combined atlases.

One weakness of the Jaccard coefficient is that it can be misleading when regions are very small or large. This because the coefficient is given a higher value when the investigated region overlaps a smaller region than when the region overlaps a large region, even though there might be more voxels that overlap with the large region than with the smaller one. In the same way the coefficient is given a lower value when the overlap is with a very large region. For this reason, the most frequent coinciding regions do not always have the largest overlaps according to the Jaccard coefficient. This might be a reason why the Jaccard coefficient for overlaps with HM or MGC background might be very small even if the overlap is quite large. Therefore, it can also be useful to use the volumetric overlap (Gerig et al., 2001).

4.4 Volumes

The mean volume ratios, mean region volumes and their CoV can be used to get an understanding of both the variation between region volumes between the different brains that are used and to see if two databases delineates the regions in the same way considering the volume. The CoV parameter made it possible to compare the variation between different regions with both large and small volumes. The volume ratio could also give an idea of the size of subdivisions relatively a whole region. The volume ratio was used because it is easily calculated and intuitive.

The observation that the HM regions had larger volumes than the MGC is explained by the fact that the HM has fewer regions. The MGC seems to have focused on smaller regions of functional importance. A factor that affected the ratios was that subdivisions were sometimes compared to a "whole" region which gave a somewhat deceptive result. The fact that the ratio between the backgrounds was almost equal to one meant that the total volume of the regions was almost equal even though they differed a lot separately. This could be because the MGC defined some very small regions that were included in bigger ones in the HM atlas and vice versa for some HM subdivisions. Even though they defined different regions that did not always coincide, the ratio implies that the atlases cover the same amount of space, without necessarily overlapping with each other proving that both atlas databases leave certain brain regions unclassified. The global comparison between the protocols could also be an indicator that the HM has larger volumes since there were 73 HM regions that matched with MGC regions while 86 MGC regions had matching HM regions. This meant that the MGC probably had more subdivisions that could be matched with a larger HM region and that 73 HM regions should roughly have the same volume as 86 MGC regions. The reason why the MGC generally had smaller and more specific regions could be because it is a commercially used atlas that is made to suit as many different customers as possible. The volume ratio between the thalamus regions was quite different when using the HM images and the MGC images respectively. This could be due to the automatization process, that causes the HM region to extend further in the MGC images and that the MGC borders are tighter in the HM images than the actual manually delineated borders were in the original MGC images. In Heckemann et al. (2006), five different discrepancies due to using the MAPER software to generate atlases were discussed. For example, one was that the software systematically included or excluded voxels at the boundaries of the label but still preserved its shape.

This could be an explanation to why the volume ratio differed but the Jaccard coefficient was almost the same for the thalamus regions.

The large difference between volumes of the HM defined and MGC defined hippocampus was unexpected, even though they had quite different borders according to the protocols. But the small absolute size of the hippocampus means that it will be more sensitive to discrepancies than larger regions. The visualization also showed that the MGC region extended further than the HM region in almost every direction. Even though the borders were close, this adds up to a large discrepancy in the total volume.

The comparison of region volumes does not regard any regional differences and does not explain where the differences occur (Gerig et al., 2001) which is the reason why some of the regions and overlaps were plotted.

The ratio between the volumes of the regions had some variations which could be due to the difference in region volumes. It also implied that the HM and MGC regions did not always correlate in their volumes, i.e. the MGC did not delineate the region as large or small as the HM did. If the region volumes had varied the same way between the images and atlases, the volume ratio between them should have been approximately the same and therefore varied less between the images than each volume separately. The standard deviation of the volume ratio could also be affected by the number of combined atlases that had the specific overlap of regions because if just a small number of atlases had the overlap, the variation would probably be smaller than if the overlap occurred in all of the images. This could also cause that the CoV would be lower for the volume ratio than for each region volume separately. The optic chiasm had a very high ratio between the standard deviation and the mean which most likely was due to its very small volume.

4.5 Evaluation of the quantitative comparison

The results from the quantitative comparison that were compared with the qualitative predictions proved that the developed quantitative method is a working concept. The results from the quantitative analysis confirmed the most important qualitative results, both the global comparison and the qualitative predictions. The global comparison was confirmed by both the results from the calculation of the Jaccard coefficient and of the most frequent coinciding regions. The predictions were confirmed based on the results from the calculation of the Jaccard coefficient, volume ratios and the most frequent coinciding regions. The quantitative analysis also revealed additional insights that the qualitative comparison did not predict. Some overlaps only occurred in some of the combined atlases which could in no way be predicted by just doing the qualitative comparison, proving the need for the quantitative comparison. The quantitative comparison solved the issue with making predictions regarding the subdivisions and its coverage of the entire region.

The quantitative method can be applied for comparative assessments of automatic segmentation methods and when comparing different segmentations of the same brain image, for example different automatic segmentations of the same brain or the use of different segmentation protocol on the same brain image. It can also be used to better understand the detailed differences for individual regions to know, for example, which atlas that have the better coverage of a specific region or to determine when one of the atlases may be more useful than the other, depending on the purpose of the use. Based on the results, the MGC atlas is preferred over the HM atlas when for example structure–functional relationships are to be mapped. However, the HM atlas is instead preferred over the MGC atlas when

for example the insula is to be examined or investigated, since the HM protocol divides the insula into more subdivisions. But considering these differences, it is not possible to state that one atlas database is comprehensively better than the other. Since the quantitative comparison yielded almost the same information as the predictions, and in most cases additional information about which regions that actually overlapped, it can be more useful than doing the comparison between the protocols. This because making the predictions was highly time-consuming and challenging, partly due to lack of descriptions of borders in some protocols. Accordingly, the quantitative comparisons described could be more informative and easily interpreted. It could also be possible to, based on the quantitative method, find differences in the protocols without having to study the protocols for each region separately.

The quantitative comparison was done for both when the combined atlases were used separately and when they were all added together, showing that the results were almost independent of method. This meant that if one only wants an overview of, for example, which regions that overlap between two atlases the results from when adding all of the combined atlases could be a simpler and fully sufficient solution. The combined atlases were analysed separately to get an understanding of how the parameters varied between the brain images and to see if the overlaps were systematic recurrences. The results were also almost independent of which database images were used and validated that the differences found between the databases were systematic. It implied that the largest overlaps and volume ratios were actual differences in the atlases and not because of the automatization proving that the two atlas databases differed systematically.

5 Conclusion

The two atlas databases differ systematically, reflecting the differences in purpose and priorities that guided the underlying manual segmentation procedures. Considering these differences, it is not possible to state that one atlas database is comprehensively better than the other. The quantitative method developed in this project confirmed the most important predictions, however, it also revealed additional insights that the qualitative analysis could not predict. The results lead to the conclusion that the HM atlas generally have larger regions than the MGC atlas. The regions that did not have a matching region in the other protocol often coincided with MGC cerebral white matter if it was a HM region and HM background if it was an MGC region. The main difference between the protocols is that cortical regions only include the actual cortical grey matter in the MGC, whereas HM includes adjacent white matter as part of the region. The largest overlaps occurred between the regions that had a matching region and the variation of the Jaccard coefficient between the brain images was quite small for these regions.

The proposed qualitative and quantitative method in this study provides a way to compare atlases and segmentations. The result was almost independent of which images were used for regions which had small standard deviations of the parameters which imply that the method is consistent and reliable. The proposed method is however not able to fully distinguish which discrepancies arise from the automatization process and which are actual differences. The visualization of the overlaps, described in the qualitative comparison, is an essential part of the comparison, given that the location of the overlap is sought after.

6 Acknowledgements

I would like to thank my supervisor, Rolf A. Heckemann for great inspiration and guidance during this project. I would also like to thank Alexander Hammers for help with finding the manual delineation protocols for regions 50–83. Last but not least, thanks to my family and friends for their endless support and patience.

7 Reference list

Bohland, J. W., Bokil, H., Allen, C. B. & Mitra, P. P., 2009. The Brain Atlas Concordance Problem: Quantitative Comparison of Anatomical Parcellations. *PLoS ONE*, 29 September, 4(9).

Faillenot, I., Heckemann, R. A., Frot, M. & Hammers, A., 2017. Macroanatomy and 3D probabilistic atlas of the human brain. *NeuroImage*, Issue 150, pp. 88 - 98.

Gerig, G., Jomier, M. & Chakos, M., 2001. *Valmet: A new validation tool for assesing and improving 3D object segmentation.* Berlin, Heidelberg, Springer Link, pp. 516 - 523.

Gousias, I. S. et al., 2008. Automatic segmentation of brain MRIs of 2-year-olds into 83 regions of interest. *NeuroImage*, Issue 40, pp. 672 - 684.

Hammers, A. et al., 2003. Three-Dimension Maximum Probability Atlas of the Human Brain, With Particular Reference to the Temporal Lobe. *Human Brain Mapping*, Volume 19, pp. 224-247.

Heckemann, R. A. et al., 2006. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage*, Volume 33, pp. 115 - 126.

Heckemann, R. A. et al., 2011. Automatic morphometry in Alzheimer's disease and mild cognitive impairment. *Neuroimage*, 15 June, 56(4), pp. 2024-2037.

Heckemann, R. A. et al., 2010. Improving intersubject image registration using tissue-class information benefits robustness and accuracy of multi-atlas based anatomical segmentation. *Neuroimage*, 15 May, 51(1), pp. 221-227.

Jaccard, P., 1912. The distribution of the flora in the alpine zone. *The new phytologist*, 29 February, XI(2).

Marcus, D. S., Wang, T. H., Parker, J. & Csernansky, J. G., 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*, September, 19(9), pp. 1498 - 1507.

Muschelli, J., 2020. *RDocumentation.* [Online]
Available at: https://www.rdocumentation.org/packages/neurobase
[Accessed 20 April 2020].

Neuromorphometrics Inc., 2005. *Segmentation: Hippocampus.* [Online]
Available at: http://neuromorphometrics.com/Seg/html/segmentation/hippocampus.html
[Accessed 10 February 2020].

Neuromorphometrics Inc., 2005. *Segmentation: Thalamus.* [Online]
Available at: http://neuromorphometrics.com/Seg/html/segmentation/thalamus.html
[Accessed 12 February 2020].

Neuromorphometrics Inc., 2018. *Detailed Description of Labeled Scans.* [Online]
Available at: http://www.neuromorphometrics.com/wp-content/uploads/2018/12/2018-10-02_DetailedDescriptionOfLabeledScans_AllScans.pdf
[Accessed 10 February 2020].

Nurbaya Yaakub, S. et al., 2020. On brain atlas choice and automatic segmentation methods: a comparison of MAPER and FreeSurfer using three atlas databases. *Sci Rep*, 18 February, Volume 10.

Ourselin, S., 2014. *Biomedical Image Registration.* London: Springer.

Tourville, J., Carper, R. & Salamon, G., 2010. *Cortical Parcellation Protocol*. [Online] Available at: http://neuromorphometrics.com/ParcellationProtocol_2010-04-05.PDF [Accessed 15 February 2020].

Wild, H. M., Heckemann, R. A., Studholme, C. & Hammers, A., 2017. Gyri of the human parietal lobe: Volumes, spatial extends, automatic labelling, and probabilistic atlases. *PLoS ONE*, 28 August, 12(8)(e0180866).

8 Appendix

The result from the global comparison.

Table 5: Result from global comparison between protocols for the MGC regions.

MGC region	MGC region name	HM region name
0	Background	Background
4	3rd Ventricle	Third ventricle
23	Right Accumbens Area	nucleus accumbens R
30	Left Accumbens Area	nucleus accumbens L
31	Right Amygdala	TL amygdala R
32	Left Amygdala	TL amygdala L
35	Brain Stem	brainstem excluding substantia nigra
36	Right Caudate	caudate nucleus R
37	Left Caudate	caudate nucleus L
38	Right Cerebellum Exterior	cerebellum R
39	Left Cerebellum Exterior	cerebellum L
40	Right Cerebellum White Matter	cerebellum R
41	Left Cerebellum White Matter	cerebellum L
47	Right Hippocampus	TL hippocampus R
48	Left Hippocampus	TL hippocampus L
49	Right Inf Lat Vent	Lateral ventricle temporal horn R
50	Left Inf Lat Vent	Lateral ventricle temporal horn L
51	Right Lateral Ventricle	Lateral ventricle excluding temporal horn R
52	Left Lateral Ventricle	Lateral ventricle excluding temporal horn L
55	Right Pallidum	pallidum R
56	Left Pallidum	pallidum L
57	Right Putamen	putamen R
58	Left Putamen	putamen L
59	Right Thalamus Proper	thalamus R
60	Left Thalamus Proper	thalamus L
71	Cerebellar Vermal Lobules I-V	cerebellum R
72	Cerebellar Vermal Lobules VI-VII	cerebellum R
73	Cerebellar Vermal Lobules VIII-X	cerebellum R
100	Right ACgG anterior cingulate gyrus	CG anterior cingulate gyrus R
101	Left ACgG anterior cingulate gyrus	CG anterior cingulate gyrus L
104	Right AOrG anterior orbital gyrus	FL anterior orbital gyrus R
105	Left AOrG anterior orbital gyrus	FL anterior orbital gyrus L
106	Right AnG angular gyrus	PL angular gyrus R
107	Left AnG angular gyrus	PL angular gyrus L
114	Right Cun cuneus	OL cuneus R

115	Left Cun cuneus	OL cuneus L
122	Right FuG fusiform gyrus	TL fusiform gyrus R
123	Left FuG fusiform gyrus	TL fusiform gyrus L
124	Right GRe gyrus rectus	FL straight gyrus R
125	Left GRe gyrus rectus	FL straight gyrus L
132	Right ITG inferior temporal gyrus	TL middle and inferior temporal gyrus R
133	Left ITG inferior temporal gyrus	TL middle and inferior temporal gyrus L
134	Right LiG lingual gyrus	OL lingual gyrus R
135	Left LiG lingual gyrus	OL lingual gyrus L
136	Right LOrG lateral orbital gyrus	FL lateral orbital gyrus R
137	Left LOrG lateral orbital gyrus	FL lateral orbital gyrus L
138	Right MCgG middle cingulate gyrus	CG posterior cingulate gyrus R
139	Left MCgG middle cingulate gyrus	CG posterior cingulate gyrus L
142	Right MFG middle frontal gyrus	FL middle frontal gyrus R
143	Left MFG middle frontal gyrus	FL middle frontal gyrus L
146	Right MOrG medial orbital gyrus	FL medial orbital gyrus R
147	Left MOrG medial orbital gyrus	FL medial orbital gyrus L
148	Right MPoG postcentral gyrus medial segment	PL postcentral gyrus R
149	Left MPoG postcentral gyrus medial segment	PL postcentral gyrus L
150	Right MPrG precentral gyrus medial segment	FL precentral gyrus R
151	Left MPrG precentral gyrus medial segment	FL precentral gyrus L
152	Right MSFG superior frontal gyrus medial segment	FL superior frontal gyrus R
153	Left MSFG superior frontal gyrus medial segment	FL superior frontal gyrus L
154	Right MTG middle temporal gyrus	TL middle and inferior temporal gyrus R
155	Left MTG middle temporal gyrus	TL middle and inferior temporal gyrus L
160	Right OFuG occipital fusiform gyrus	TL fusiform gyrus R
161	Left OFuG occipital fusiform gyrus	TL fusiform gyrus L
162	Right OpIFG opercular part of the inferior frontal gyrus	FL inferior frontal gyrus R
163	Left OpIFG opercular part of the inferior frontal gyrus	FL inferior frontal gyrus L
164	Right OrIFG orbital part of the inferior frontal gyrus	FL inferior frontal gyrus R
165	Left OrIFG orbital part of the inferior frontal gyrus	FL inferior frontal gyrus L
166	Right PCgG posterior cingulate gyrus	CG posterior cingulate gyrus R
167	Left PCgG posterior cingulate gyrus	CG posterior cingulate gyrus L
170	Right PHG parahippocampal gyrus	TL parahippocampal and ambient gyrus R
171	Left PHG parahippocampal gyrus	TL parahippocampal and ambient gyrus L
176	Right PoG postcentral gyrus	PL postcentral gyrus R
177	Left PoG postcentral gyrus	PL postcentral gyrus L

178	Right POrG posterior orbital gyrus	FL posterior orbital gyrus R
179	Left POrG posterior orbital gyrus	FL posterior orbital gyrus L
182	Right PrG precentral gyrus	FL precentral gyrus R
183	Left PrG precentral gyrus	FL precentral gyrus L
186	Right SCA subcallosal area	FL subcallosal area R
187	Left SCA subcallosal area	FL subcallosal area L
190	Right SFG superior frontal gyrus	FL superior frontal gyrus R
191	Left SFG superior frontal gyrus	FL superior frontal gyrus L
194	Right SMG supramarginal gyrus	PL supramarginal gyrus R
195	Left SMG supramarginal gyrus	PL supramarginal gyrus L
198	Right SPL superior parietal lobule	PL superior parietal gyrus R
199	Left SPL superior parietal lobule	PL superior parietal gyrus L
204	Right TrIFG triangular part of the inferior frontal gyrus	FL inferior frontal gyrus R
205	Left TrIFG triangular part of the inferior frontal gyrus	FL inferior frontal gyrus L

Table 6: Result from global comparison between protocols for the HM regions.

HM region	**HM region name**	**MGC region name**
0	Background	Background
1	TL hippocampus R	Right Hippocampus
2	TL hippocampus L	Left Hippocampus
3	TL amygdala R	Right Amygdala
4	TL amygdala L	Left Amygdala
9	TL parahippocampal and ambient gyrus R	Right PHG parahippocampal gyrus
10	TL parahippocampal and ambient gyrus L	Left PHG parahippocampal gyrus
11	TL superior temporal gyrus middle part R	Right STG superior temporal gyrus
12	TL superior temporal gyrus middle part L	Left STG superior temporal gyrus
15	TL fusiform gyrus R	Right FuG fusiform gyrus
16	TL fusiform gyrus L	Left FuG fusiform gyrus
19	brainstem excluding substantia nigra	Brain Stem
20	insula posterior long gyrus L	Left PIns posterior insula
21	insula posterior long gyrus R	Right PIns posterior insula
24	CG anterior cingulate gyrus L	Left ACgG anterior cingulate gyrus
25	CG anterior cingulate gyrus R	Right ACgG anterior cingulate gyrus
26	CG posterior cingulate gyrus L	Left PCgG posterior cingulate gyrus
27	CG posterior cingulate gyrus R	Right PCgG posterior cingulate gyrus
28	FL middle frontal gyrus L	Left MFG middle frontal gyrus
29	FL middle frontal gyrus R	Right MFG middle frontal gyrus
32	PL angular gyrus L	Left AnG angular gyrus
33	PL angular gyrus R	Right AnG angular gyrus

34	caudate nucleus L	Left Caudate
35	caudate nucleus R	Right Caudate
36	nucleus accumbens L	Left Accumbens Area
37	nucleus accumbens R	Right Accumbens Area
38	putamen L	Left Putamen
39	putamen R	Right Putamen
40	thalamus L	Left Thalamus Proper
41	thalamus R	Right Thalamus Proper
42	pallidum L	Left Pallidum
43	pallidum R	Right Pallidum
45	Lateral ventricle excluding temporal horn R	Right Lateral Ventricle
46	Lateral ventricle excluding temporal horn L	Left Lateral Ventricle
47	Lateral ventricle temporal horn R	Right Inf Lat Vent
48	Lateral ventricle temporal horn L	Left Inf Lat Vent
49	Third ventricle	3rd Ventricle
50	FL precentral gyrus L	Left PrG precentral gyrus
51	FL precentral gyrus R	Right PrG precentral gyrus
52	FL straight gyrus L	Left GRe gyrus rectus
53	FL straight gyrus R	Right GRe gyrus rectus
54	FL anterior orbital gyrus L	Left AOrG anterior orbital gyrus
55	FL anterior orbital gyrus R	Right AOrG anterior orbital gyrus
58	FL superior frontal gyrus L	Left SFG superior frontal gyrus
59	FL superior frontal gyrus R	Right SFG superior frontal gyrus
60	PL postcentral gyrus L	Left PoG postcentral gyrus
61	PL postcentral gyrus R	Right PoG postcentral gyrus
62	PL superior parietal gyrus L	Left SPL superior parietal lobule
63	PL superior parietal gyrus R	Right SPL superior parietal lobule
64	OL lingual gyrus L	Left LiG lingual gyrus
65	OL lingual gyrus R	Right LiG lingual gyrus
66	OL cuneus L	Left Cun cuneus
67	OL cuneus R	Right Cun cuneus
68	FL medial orbital gyrus L	Left MOrG medial orbital gyrus
69	FL medial orbital gyrus R	Right MOrG medial orbital gyrus
70	FL lateral orbital gyrus L	Left LOrG lateral orbital gyrus
71	FL lateral orbital gyrus R	Right LOrG lateral orbital gyrus
72	FL posterior orbital gyrus L	Left POrG posterior orbital gyrus
73	FL posterior orbital gyrus R	Right POrG posterior orbital gyrus
78	FL subcallosal area L	Left SCA subcallosal area
79	FL subcallosal area R	Right SCA subcallosal area
82	TL superior temporal gyrus anterior part L	Left STG superior temporal gyrus
83	TL superior temporal gyrus anterior part R	Right STG superior temporal gyrus

84	PL supramarginal gyrus L	Right SMG supramarginal gyrus
85	PL supramarginal gyrus R	Left SMG supramarginal gyrus
86	insula anterior short gyrus L	Left AIns anterior insula
87	insula anterior short gyrus R	Right AIns anterior insula
90	insula posterior short gyrus L	Left PIns posterior insula
91	insula posterior short gyrus R	Right PIns posterior insula
92	insula anterior pole L	Left AIns anterior insula
93	insula anterior pole R	Right AIns anterior insula
94	insula anterior long gyrus L	Left AIns anterior insula
95	insula anterior long gyrus R	Right AIns anterior insula