Parcellation of the cortex using Surface-based Melbourne Children's Regional Infant Brain atlases (M-CRIB-S)

Chris L. Adamson[1], Bonnie Alexander[1], Gareth Ball[1], Richard Beare[1,2], Jeanie L.Y. Cheong[1,3,4], Alicia J. Spittle[1,3,5], Lex W. Doyle[1,3,4,6], Peter J. Anderson[1 6,7], Marc L. Seal[1,6] Deanne K. Thompson[1,6,8]

[1]Murdoch Children's Research Institute, Melbourne, Victoria, Australia

[2]Department of Medicine, Monash University, Melbourne, Australia

[3]Neonatal services, The Royal Women's Hospital, Melbourne, Australia

[4]Department of Obstetrics and Gynaecology, The University of Melbourne, Melbourne, Australia

[5]Department of Physiotherapy, The University of Melbourne, Melbourne, Australia

[6]Department of Paediatrics, The University of Melbourne, Melbourne, Australia

[7]Monash Institute of Cognitive and Clinical Neurosciences, Monash University, Melbourne, Australia

[8]Florey Department of Neuroscience and Mental Health, The University of Melbourne, Melbourne, Australia

Corresponding author

Associate Professor Deanne K. Thompson, Murdoch Children's Research Institute, Royal Children's Hospital, Flemington Road, Parkville, Victoria, 3052, Australia. Telephone +61 3 9936 6708. Fax +61 3 9348 1391. Email deanne.thompson@mcri.edu.au

Abbreviations

BSTS: Banks of the Superior Temporal Sulcus, CAC: Caudal anterior cingulate, CMF: Caudal middle frontal, CUN: Cuneus, dHCP: developing Human Connectome Project, DK: Desikan-Killiany , DKT: Desikan-Killiany-Tourville, ENT: Entorhinal, FP: Frontal Pole, FUS: Fusiform, INFP: Inferior parietal, INS: Insula, ISTC: Isthmus cingulate, IT: Inferior temporal, L+U: Labelled+Unlabelled, LIN: Lingual, LOCC: Lateral occipital, LOO: Leave One Out, LORB: Lateral orbito-frontal, M-CRIB: Melbourne Children's Regional Infant Brain, M-CRIB-S: Surface-based Melbourne Children's Regional Infant

Brain atlas, MORB: Medial orbito-frontal, MR: Magnetic Resonance Imaging, MT: Middle temporal, PARC: Paracentral lobule, PARH: Parahippocampal, PC: Posterior cingulate, PCAL: Pericalcarine, PCUN: Precuneus, POPE: Pars opercularis, PORB: Pars orbitalis, PREC: Precentral, PTRI: Pars triangularis, RAC: Rostral anterior cingulate, RMF: Rostral middle frontal, SF: Superior frontal, SMAR: Supramarginal, SP: Superior parietal, ST: Superior temporal, TP: Temporal Pole, TT: Transverse temporal, UNC: University of North Carolina.

Key Words

Neonate, baby, segmentation, magnetic resonance imaging, cortical, sulcus, gyrus

Highlights

- Surface-based neonatal brain atlases have been made available
- An automated infant brain parcellation pipeline has been released
- This is an accurate and validated technique to obtain infant cortical regions
- The parcellation scheme matches commonly used adult atlases
- These tools enable assessment of cortical development across the lifespan

Abstract

Background

Longitudinal studies of cortical morphology are best facilitated by parcellation schemes that are compatible across all life stages. Until recently, neonatal atlases derived strictly from neonatal data, rather than warping brain labels from an adult atlas, did not exist. The Melbourne Children's Regional Infant Brain (M-CRIB) and M-CRIB 2.0 atlases now provide voxel-based parcellations of the cerebral cortex within T_2-weighted neonatal images. These atlases are compatible with the Desikan-Killiany (DK) and the Desikan-Killiany-Tourville (DKT) cortical labelling schemes commonly used for older children and adults. However, there is still a need for a surface-based approach for parcellating neonatal images using these atlases.

Aims

We aimed to introduce surface-based versions of the M-CRIB and M-CRIB 2.0 atlases, termed M-CRIB-S(DK) and M-CRIB-S(DKT), along with a pipeline for automated parcellation with FreeSurfer tools. We also aimed to evaluate the automated parcellation accuracy of our M-CRIB-S atlases using cross-validation with manually labelled data; and to evaluate accuracy of the proposed automated parcellation pipeline for the M-CRIB-S labels against the publicly available surface-based University of North Carolina (UNC) 4D neonatal atlas, which has labels derived from the DK adult atlas.

Methods

Using datasets of 10 M-CRIB or M-CRIB 2.0 manually labelled ground truth images and 48 unlabelled T_2-weighted magnetic resonance images (MRI) of healthy neonates, cortical surfaces were extracted using the *Deformable* module within *MIRTK*. Cortical regions from the labelled images were encoded into a spherical template space using *FreeSurfer* tools wherein M-CRIB-S(DK) and M-CRIB-S(DKT) atlases were constructed. Automatic parcellation of the labelled and unlabelled (L+U) images (n=58) using the M-CRIB-S(DK) and M-CRIB-S(DKT) atlases were computed with Bayesian labelling using *FreeSurfer* tools. Accuracy was assessed by comparison with manually labelled ground truth data (n=10) using Dice coefficients and within a Leave-One-Out (LOO) cross-validation framework.

Agreement of the UNC atlas parcellations with manual M-CRIB labelling was also evaluated using Dice coefficients.

Results

When comparing automated parcellations from the M-CRIB-S atlases to manual labels, disagreements were mostly closely confined to region boundaries, suggesting high overlap overall. Quantitatively, average regional Dice coefficients were 0.88 (L+U) and 0.83 (LOO) for the M-CRIB 2.0 labels and 0.87 (L+U), 0.80 (LOO) for the M-CRIB labels, and 0.67 for the UNC atlas.

Conclusions

The M-CRIB-S atlases and automatic pipeline described in this paper allow infant cortical surfaces to be accurately parcellated according to the DK or DKT parcellation schemes. This will help facilitate surface-based investigations of brain function at the neonatal time point, and potentially longitudinally across the lifespan. The use of M-CRIB-S outperformed a technique that relies on an adult-derived atlas. The atlases, averaged white, inflated, pial and spherical surfaces, along with customised scripts for segmentation, cortical surface extraction and parcellation, are available for public download.

Introduction

The delineation of cortical areas on magnetic resonance images (MRI) to provide a map of the brain based on structural or functional organisation, cytoarchitecture or shared connectivity, is often considered a prerequisite for beginning to understand the complexities of the human brain (Amunts and Zilles, 2015; Brett et al., 2002; Glasser et al., 2016). In particular, understanding *development* of the brain is reliant on accurate parcellation of the cerebral cortex at one of the most critical times of development, around birth. *FreeSurfer* (Fischl, 2012) is a commonly used cortical parcellation pipeline used for children and adults, and its labelling schemes include the Desikan-Killiany (DK) (Desikan et al., 2006) and Desikan-Killiany-Tourville (DKT) (Klein and Tourville, 2012) adult atlases. However, tools tuned for adult brains, such as the adult T_1-based templates and atlases provided in *FreeSurfer*, are not directly applicable to neonatal brain images. The optimal MRI sequences for the grey and white matter contrast required to identify cortical surfaces differ by age. While T_1-weighting is optimal for adult brains, T_2-weighting contrasts are optimal for neonatal brains. Consequently, specialised algorithms for neonatal data are required in order to contend with the differences in tissue signal intensities in grey and white matter compared to adults (Beare et al., 2016; Makropoulos et al., 2016). Thus, brain atlases and image segmentation and parcellation tools specific for infants are required.

Methods for cortical parcellation of infant brain images have focused on warping standardised cortical atlases from adult brains (Desikan et al., 2006; Klein and Tourville, 2012; Tzourio-Mazoyer et al., 2002) to infant brains (Shi et al., 2011; Wu et al., 2018). In one example, volumetric labelling of infant brains was achieved by projecting the adult Automated Anatomical Labelling (AAL) (Tzourio-Mazoyer et al., 2002) template onto infant brain images (Shi et al., 2011) which required warping infant brains to adult templates. Another example is the University of North Carolina (UNC) 4D neonatal atlas, which contains a projection of the DK adult atlas (Desikan et al., 2006) onto neonatal cortical surfaces (Wu et al., 2018) for gestational ages between 39 to 44 weeks inclusive (Li et al.,

2015). However, labelling a neonatal brain image using adult-derived atlases is problematic, due to the inherent difference in anatomy and tissue composition between infant and adult brains (Richards et al., 2016; Richards and Xie, 2015).

Recently we introduced neonatal-specific voxel-wise brain atlases compliant with the widely-used adult atlases from *FreeSurfer*. The Melbourne Children's Regional Infant Brain (M-CRIB) atlas (Alexander et al., 2017) comprises 34 cortical regions per hemisphere, compatible with the DK labelling scheme (Desikan et al., 2006), and the M-CRIB 2.0 atlas (Alexander et al., 2018) contains 31 cortical labels replicating the DKT (Klein and Tourville, 2012) labelling scheme. Each of the M-CRIB atlases comprise 10 manually-segmented neonatal brains, providing accurate coverage of the neonatal brain with ground truth manual segmentations and clear areal boundary definitions.

While accurate labelling can be achieved using voxel-based parcellation schemes (Alexander et al., 2018; Alexander et al., 2017; Makropoulos et al., 2018), surface-based registration methods lead to significantly improved alignment of cortical landmarks, including cortical folds, and therefore more accurate placement of areal boundaries (Coalson et al., 2018; Ghosh et al., 2010). Until now, a surface-based cortical parcellation method specific to the infant brain has not been available.

The Developing Human Connectome Project (dHCP) has recently provided a pipeline for segmentation and cortical extraction for T_2-weighted images of neonatal brains (Bozek et al., 2018; Makropoulos et al., 2018). This process segments brain tissue into cerebral and cerebellar grey and white matter, and various subcortical grey matter structures, before extracting an inner and outer cortical surface which is automatically partitioned into lobes. These existing cortical surface extraction tools can be expanded upon to incorporate a neonatal brain atlas to provide accurate surface-based parcellation. Measures such as cortical thickness, surface area, and curvature for each cortical sub-division can subsequently be derived for infant MRI scans.

This study aimed to expand on available cortical extraction approaches to present a surface-based representation of the DK-compatible M-CRIB and DKT-compatible M-CRIB 2.0 atlases based on gold-standard, manually labelled neonatal datasets. Additionally, we aimed to provide neonatal average surface templates and surface-based cortical atlases based on the M-CRIB 2.0 and M-CRIB parcellation schemes that are compatible with FreeSurfer and the dHCP pipelines (Fischl, 2012; Fischl et al., 1999; Fischl et al., 2004), along with companion scripts to perform cortical surface extraction, surface registration and labelling of a novel neonatal T_2-weighted brain images using the atlases. These tools will, for the first time, enable automated parcellation of the neonatal cortex that is compatible with other available pipelines and therefore able to be employed for longitudinal studies.

Methods

Participants

A total of 58 term-born (≥37 weeks' gestation), healthy neonates (40.2-44.9 weeks post-menstrual age (PMA) at scan, M = 42.4, SD = 1.2, 26 female) were scanned as control subjects as part of preterm birth studies (Spittle et al., 2014; Walsh et al., 2014). Criteria for a subject being healthy were no admissions to a neonatal intensive care or special care unit, resuscitation at birth not required, birthweight more than 2.5 kg and no evidence of congenital conditions known to affect development and growth. Ethical approval for the studies was obtained from the Human Research Ethics Committees of the Royal Women's Hospital and the Royal Children's Hospital, Melbourne. Written informed consent was obtained from all parents. Data that exhibited excessive movement or other corrupting artefacts were excluded. This cohort was subdivided into the following two subsets: *labelled* and *unlabelled* subsets. The *labelled* set comprised the ten subjects (40.29 – 43.00 weeks' PMA at scan, M = 41.71, SD = 1.31, 4 female) of the M-CRIB atlas that had been previously randomly selected from this cohort (Alexander et al., 2018; Alexander et al., 2017). The *unlabelled*

subset consisted of the remaining 48 subjects (40.2 – 44.9 weeks' PMA at scan, *M* = 42.6, *SD* = 1.3, 22 female).

MRI acquisition

Subjects were scanned at the Royal Children's Hospital, Melbourne, Australia, on a 3T Siemens Magnetom Trio scanner during unsedated sleep. T_2-weighted images were acquired with a turbo spin echo sequence with the following parameters: 1 mm axial slices, flip angle = 120°, repetition time= 8910 ms, echo time = 152 ms, field of view = 192 × 192 mm, in-plane resolution = 1 mm^2 (zero-filled interpolated to 0.5 × 0.5 × 1 mm), matrix size = 384 × 384. All T_2-weighted images were resliced to voxel-volume-preserving size of 0.63 × 0.63 × 0.63 mm.

Image segmentation

The *unlabelled* dataset was segmented automatically using the *DrawEM* software package (Makropoulos et al., 2016; Makropoulos et al., 2014). Briefly, this technique non-linearly registered the non-labelled T_2-weighted images to multiple pre-labelled images. The non-labelled image was then segmented using label fusion. The output of *DrawEM* included voxel-wise segmentations of cerebral white and grey matter (including lobar sub-divisions), cerebellum and various subcortical grey matter structures.

The labelled dataset segmentations were taken directly from the M-CRIB and M-CRIB 2.0 atlases. Each segmented image comprised manually traced cerebral white and grey matter, cerebellum, basal ganglia and thalamus, cortical, ventricular and other subcortical labels. Tracing protocols for the cortical (Alexander et al., 2018; Alexander et al., 2017) and subcortical (Loh et al., 2016) segmentations have been previously described. Translation of label numbers was performed for compatibility with *DrawEM*.

Surface extraction

The *Deformable* module (Bozek et al., 2018; Schuh et al., 2017) of MIRTK (https://github.com/BioMedIA/MIRTK) was employed to extract the inner and outer boundaries of each hemisphere of the cerebral cortices for both labelled and unlabelled datasets. For each subject, this algorithm required the T_2-weighted image and the corresponding segmentation containing cerebral white/grey, subcortical grey and cerebellum structure labels.

Curvature template generation

Surface templates are constructed using the curvature-based spherical mapping, alignment and averaging method as previously described (Fischl, 2012; Fischl et al., 1999). Briefly, to construct a surface template from healthy neonatal subjects we initially used the white matter surface mesh of one subject chosen randomly to be the initial target. An iterative procedure of aligning spherical surfaces from both the *labelled* and *unlabelled* datasets to the current target, followed by creation of a new target, was performed. The alignment of a subject's spherical surface involved distorting coordinates to those of the template to maximise similarity of curvatures and minimise geometric distortions to the original spherical surface. The use of local curvatures and sulcal depth to drive registration meant that corresponding sulci and gyri were aligned. The final template curvature map was created by averaging all aligned curvature maps.

The spherical mapping of each white matter surface onto a common spherical space meant that any given point in template space could be mapped to a point on each subject's white matter surface, and those points were in correspondence across subjects. This enabled average white, pial and inflated surfaces to be constructed using the *FreeSurfer* tool *mris_make_average_surface*, by resampling surfaces onto a common icosahedron. Figure 1 displays voxel-based segmentation and estimated white, inflated and spherical cortical surfaces for one subject from the labelled and unlabelled sets.

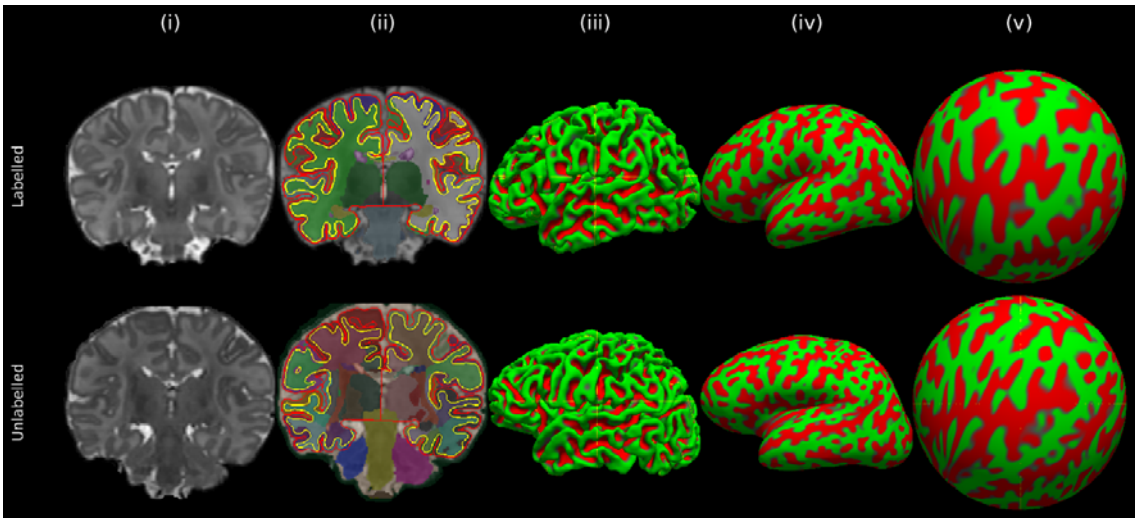


Figure 1: Exemplary segmentation, surface extraction and spherical projection for one subject from each of the labelled (top panels) and unlabelled sets (bottom panels). (i) Original T_2-weighted images, (ii) segmentations according to the M-CRIB 2.0 (top) and *DrawEM* (bottom) techniques, the yellow and red contours show the inner and outer cortical surface in the current slice, (iii) three-dimensional view of the white matter surfaces from (ii), (iv) inflated surfaces, (v) spherical projection of the white matter surfaces from (iii). The colours in (iii), (iv) and (v) denote curvature sign with red and green representing sulci and gyri, respectively.

Parcellation training set construction

Parcellation training sets were constructed using the *labelled* set for each M-CRIB and M-CRIB 2.0 cortical label, using the method of Fischl et al. (Fischl et al., 2004). Firstly, the voxel-based parcellation labels of the cortex were projected onto white matter surface vertices using a constrained nearest-labelled voxel search. Automated post-processing was performed on each label to remove all but the largest connected components. Spatial priors in registered spherical surface space were then constructed using the tool *mris_ca_train*.

The surface based parcellations comprising the M-CRIB and M-CRIB 2.0 training sets made up the M-CRIB-S(DK) and M-CRIB-S(DKT) atlases, respectively. M-CRIB-S data were individually checked for anatomical accuracy of label placement by one author (B.A.). For both atlases, label placement was considered to be highly accurate. In a few instances, very minor mislabelling was identified and manually corrected on the relevant surface, and corrected volumetrically in some cases for M-CRIB 2.0 data.

Applying novel images to the template

Novel T_2-weighted images may be parcellated using the M-CRIB-S atlas data using the following sequence of processing steps: 1) Apply *DrawEM* and *Deformable* to extract white and pial surfaces, 2) Perform surface inflation, spherical projection and registration to the M-CRIB-S atlas data using FreeSurfer tools, and 3) Use the automatic labelling tool *mris_ca_label* to parcellate the surfaces. A collection of scripts are provided to execute the pipeline, which can be found along with the M-CRIB-S data on the github page (https://www.github.com/DevelopmentalImagingMCRI/M-CRIB-S). This pipeline was used to perform cortical parcellation on the *unlabelled* images for validation purposes.

Classification accuracy tests

Parcellation accuracy of the automatic labelling scheme using M-CRIB-S atlases was quantified per region and per hemisphere by calculating Dice Coefficient label overlap measures with ground-truth manually derived M-CRIB labels. Two train/test methodologies were utilised: 1. An average curvature template was constructed from all labelled and unlabelled subjects and a label training set from all labelled subjects, then each labelled subject was parcellated automatically, and 2. A Leave-One-Out (LOO) cross-validation method was utilised for the *labelled* set only such that, for each of the 10 images, curvature templates and labelling training data were created from the remaining nine. The left-out subject surfaces were then registered to the spherical templates and labelled.

Comparison with another label atlas

The UNC neonatal atlas (Li et al., 2015; Wu et al., 2018) contains averaged white, inflated and spherical surfaces with accompanying DK cortical labels. Unlike a label training set comprised of multiple parcellations on a standard surface, this dataset contains a single labelling for each template surface. To compare the accuracy of the proposed method with the UNC atlas, we projected the 42-week PMA UNC atlas labels onto each manually labelled surface by non-linear spherical registration (*mris_register*). The 42-week atlas was chosen as it most closely matched the average age of the labelled subjects (41.7 weeks PMA). Dice Coefficients, indicating label overlap with manually defined M-CRIB labels, were computed.

Results

A typical subject from the M-CRIB-S(DKT) atlas is shown in Figure 2, with inflated surfaces labelled according to the manually defined labels from M-CRIB 2.0. Figure S1 shows the equivalent M-CRIB-S(DK) atlas, labelled according to manually defined M-CRIB labels.

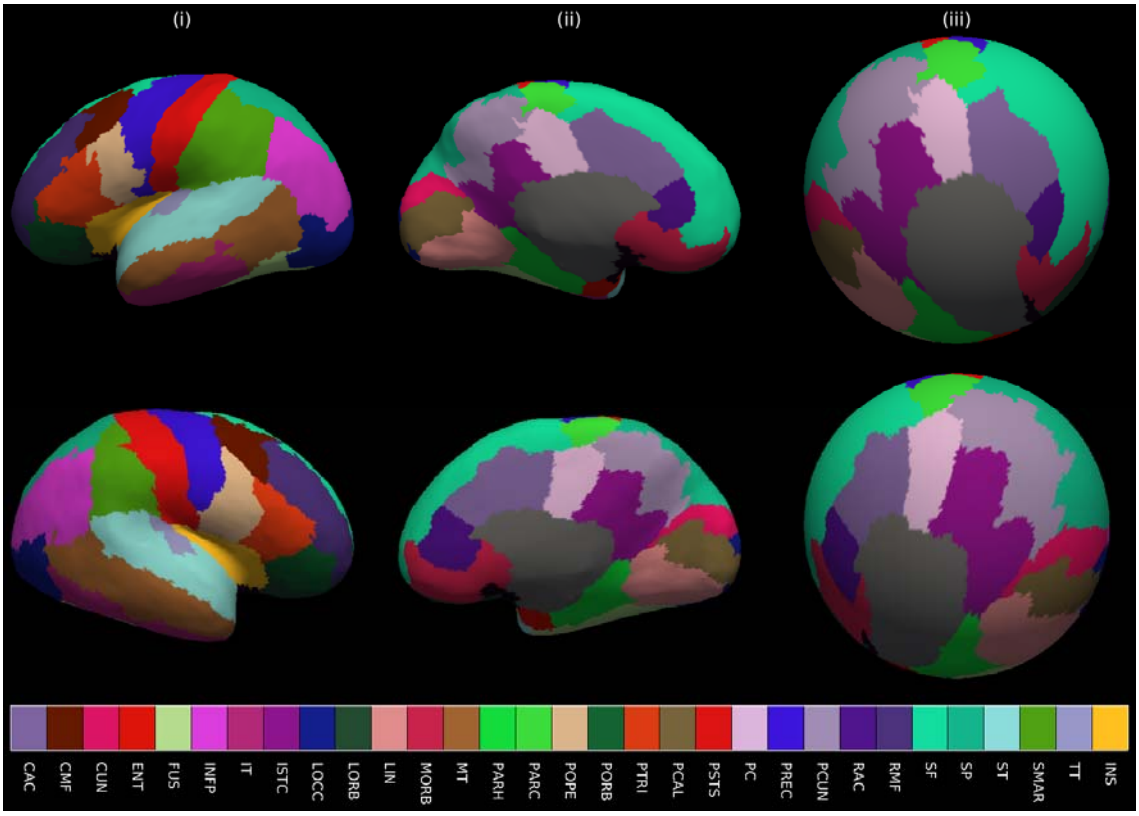


Figure 2: Exemplary manual parcellation for one subject from the labelled set using the M-CRIB 2.0 labels. (i) and (ii) show the labelling on the "inflated" surface showing left (top) and right (bottom) hemispheres. (i) Lateral aspect, (ii) medial aspect, and (iii) spherical projection with labels overlaid. CAC = caudal anterior cingulate, CMF = caudal middle frontal, CUN = cuneus, ENT = entorhinal, FUS = fusiform, INFP = inferior parietal, IT = inferior temporal, ISTC = isthmus cingulate, LOCC = lateral occipital, LORB = lateral orbitofrontal, LIN = lingual, MORB = medial orbitofrontal, MT = middle temporal, PARH = parahippocampal, PARC = paracentral, POPE = pars opercularis, PORB =pars orbitalis, PTRI = pars triangularis, PCAL = pericalcarine, PSTS = postcentral, PC = posterior cingulate, PREC = precentral, PCUN = precuneus, RAC = rostral anterior cingulate, RMF = rostral middle frontal, SF = superior frontal, SP = superior parietal, ST = superior temporal, SMAR = supramarginal, TT = transverse temporal, INS = insula.

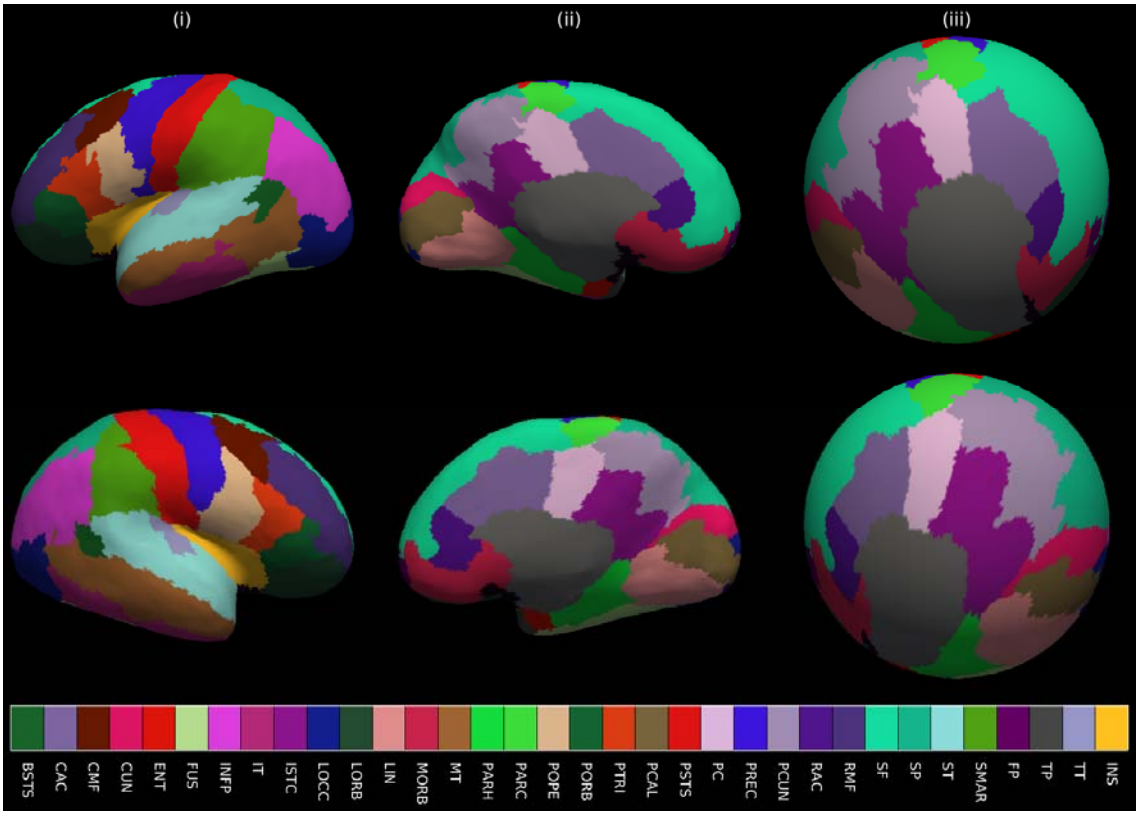


Figure S1: Exemplary manual labelling for one subject from the labelled set. (i) and (ii) show the labelling on the "inflated" surface showing left (top) and right (bottom) hemispheres. (i) Lateral aspect (iii) spherical projection with labels overlaid. BSTS = banks of the superior temporal sulcus, CAC = caudal anterior cingulate, CMF = caudal middle frontal, CUN = cuneus, ENT = entorhinal, FUS = fusiform, INFP = inferior parietal, IT = inferior temporal, ISTC = isthmus cingulate, LOCC = lateral occipital, LORB = lateral orbitofrontal, LIN = lingual, MORB = medial orbitofrontal, MT = middle temporal, PARH = parahippocampal, PARC = paracentral, POPE = pars opercularis, PORB =pars orbitalis, PTRI = pars triangularis, PCAL = pericalcarine, PSTS = postcentral, PC = posterior cingulate, PREC = precentral, PCUN = precuneus, RAC = rostral anterior cingulate, RMF = rostral middle frontal, SF = superior frontal, SP = superior parietal, ST = superior temporal, SMAR = supramarginal, FP = frontal pole, TP = temporal pole TT = transverse temporal, INS = insula.

Automatic labelling of the average white matter surface of the L+U data using the M-CRIB-S(DKT) atlas is shown in Figure 3. Figure S2 shows the equivalent automatic labelling of the average white matter surface using the M-CRIB-S(DK) atlas. These surfaces and labels may be used for display of statistical analysis results.

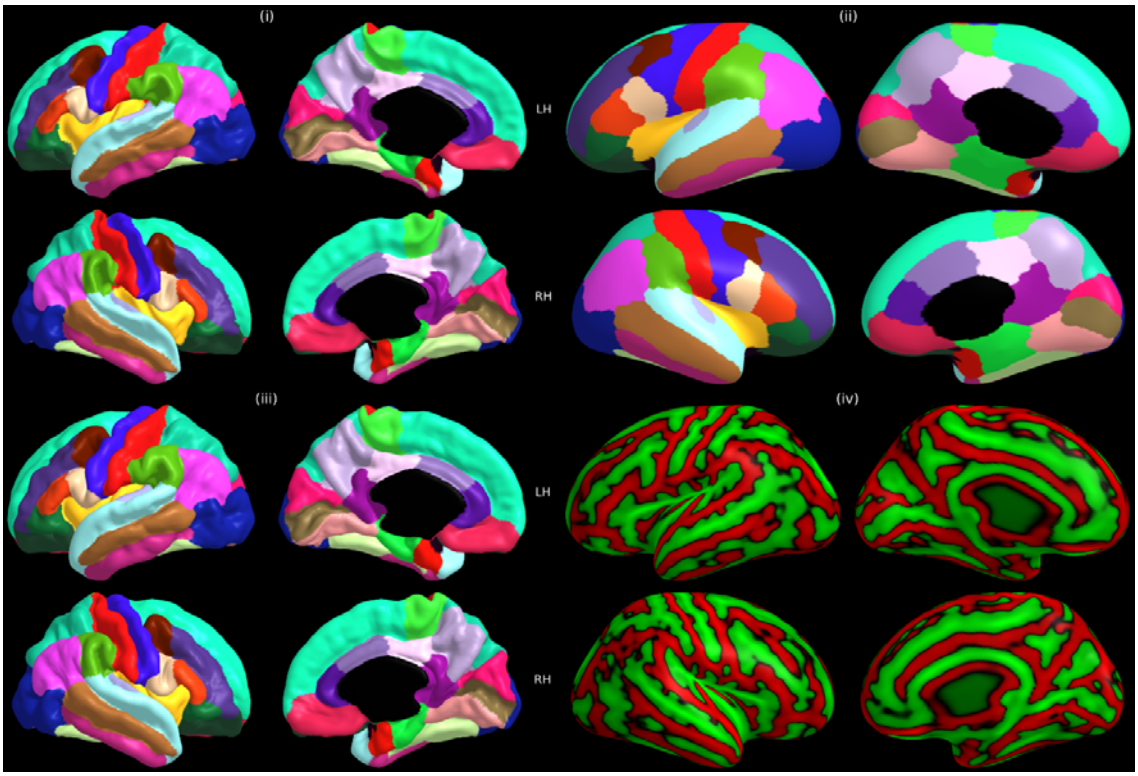


Figure 3: Average white (i), inflated (ii), and pial (iii) surfaces for L+U subjects with the vertices coloured according to their M-CRIB-S(DKT) labels . The per vertex labels are computed as a majority vote across all training parcellations in template space. The average white curvature map is shown painted onto the inflated surfaces in (iv).

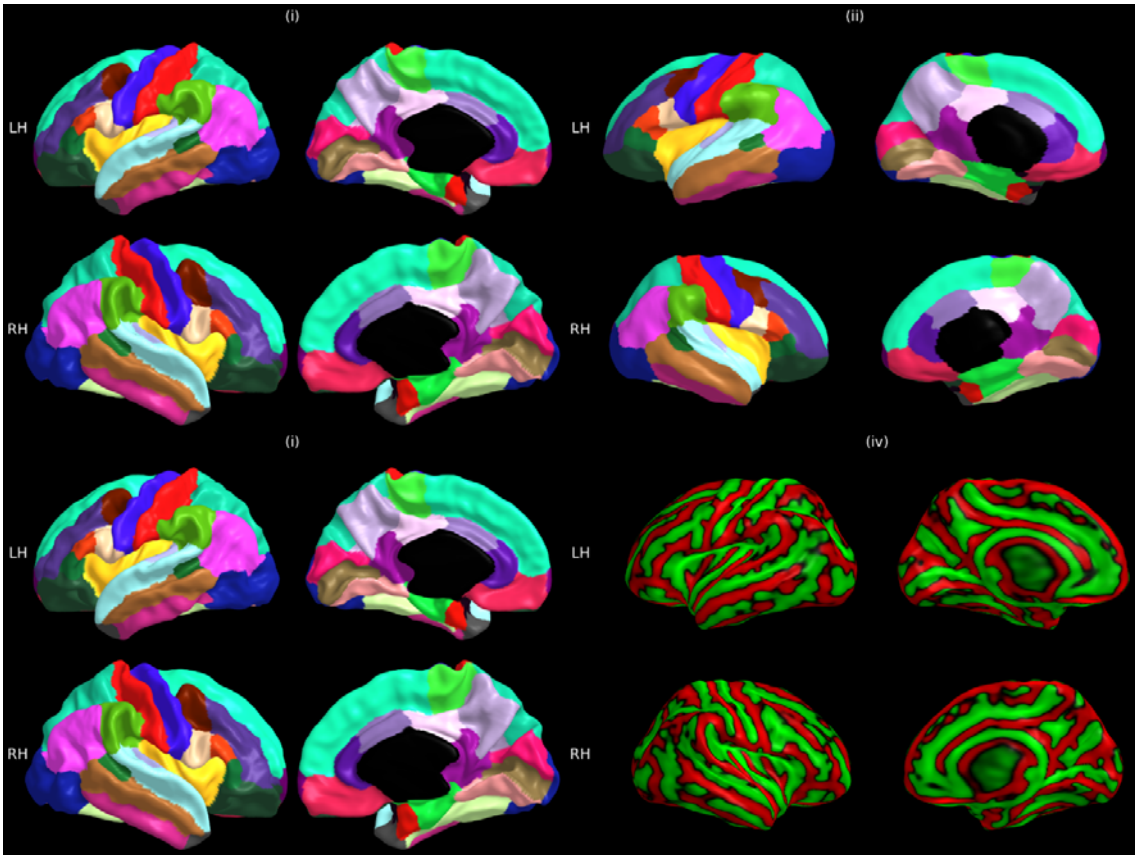


Figure S2: Average white (i), inflated (ii), and pial (iii) surfaces for L+U subjects with the vertices coloured according to their M-CRIB-S(DK) labels. The per vertex labels are computed as a majority vote across all training parcellations in template space. The average white curvature map is shown painted onto the inflated surfaces in (iv).

The dataset released in this paper is comprised of curvature templates and averaged white, inflated and pial surfaces derived from the L+U set along with the individual datasets of the M-CRIB-S(DK) and M-CRIB-S(DKT) atlases derived from the M-CRIB and M-CRIB 2.0 labelled data. For interoperability with the dHCP and UNC atlases, we also provide versions of the M-CRIB spherical template surfaces registered to the dHCP 42-week and UNC 42-week spherical template surfaces.

Figure 4 shows vertex-wise parcellation error rates across all labelled subjects for the M-CRIB-S(DKT), M-CRIB-S(DK) (including L+U and LOO curvature templates), and the UNC parcellation

schemes. Across all atlases most errors were concentrated closely around the manually defined boundaries. Thus, agreement of the estimated region locations with those manually defined was high. The UNC atlas had high mismatch rates in the inferior frontal gyrus (pars triangularis (PTRI), pars orbitalis (PORB) and pars opercularis (POPE) regions). All the M-CRIB labels had higher mismatch rates at the occipital border of the superior temporal sulcus.

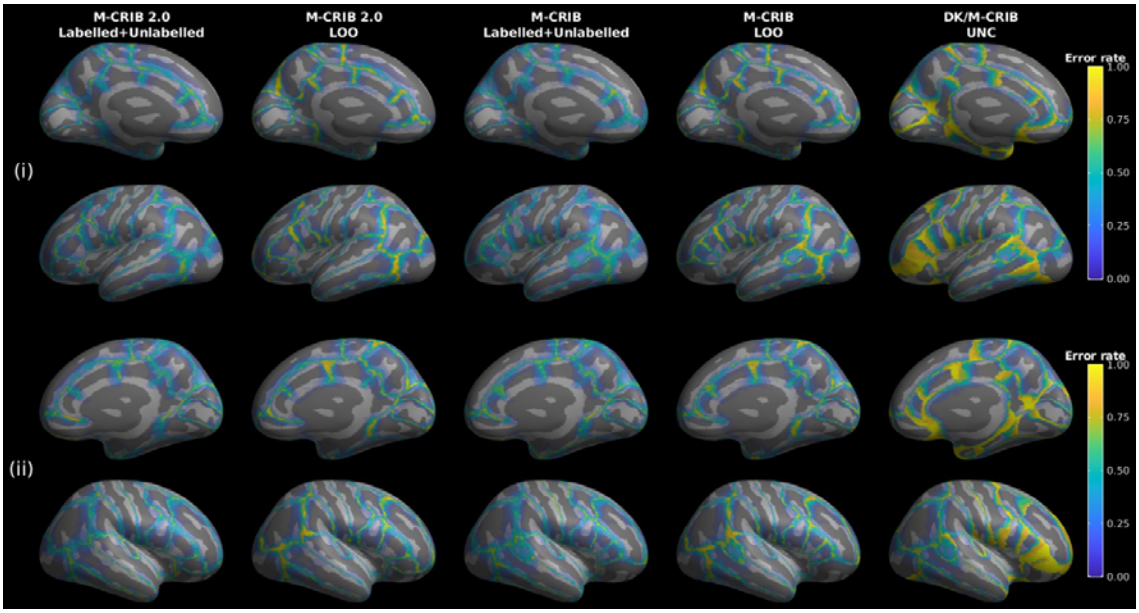


Figure 4: Inflated left (i) and right (ii) hemisphere average surfaces displaying classification error rates for both M-CRIB 2.0 and M-CRIB labelling schemes under the labelled and unlabelled (L+U) and leave one out (LOO) curvature template and training scenarios. The rightmost column displays the classification error rates for the UNC atlas using the DK labelling scheme compared to the ground truth M-CRIB labels.

Figure 5 shows quantitative regional and hemisphere-wide measures of overlap between automated and manual labelling using the M-CRIB 2.0 parcellation scheme for both the L+U and LOO curvature templates. Regional means for the L+U curvature templates ranged from 0.74-0.95 for the left hemisphere and 0.75-0.95 for the right hemisphere and the LOO curvature templates ranged from

0.83-0.97 (left) and 0.83-0.97 (right). Universally the insula (INS) had the highest mean accuracy. The lowest mean accuracies were in the following regions: left PORB, right cuneus (CUN) (L+U) and left POPE, right posterior cingulate (PC) (LOO). Low-accuracy outliers were observed in the left caudal anterior cingulate (CAC), left POPE, left PTRI, left caudal middle frontal (CMF), left fusiform (FUS), right posterior cingulate (PC), right caudal middle frontal (CMF) and right inferior parietal (INFP). Whole-hemisphere mean accuracies were 0.90 for both left and right hemispheres in L+U and 0.84 (left) and 0.85 (right) for LOO. Average regional Dice coefficients were 0.90 (L+U) and 0.83 (LOO).

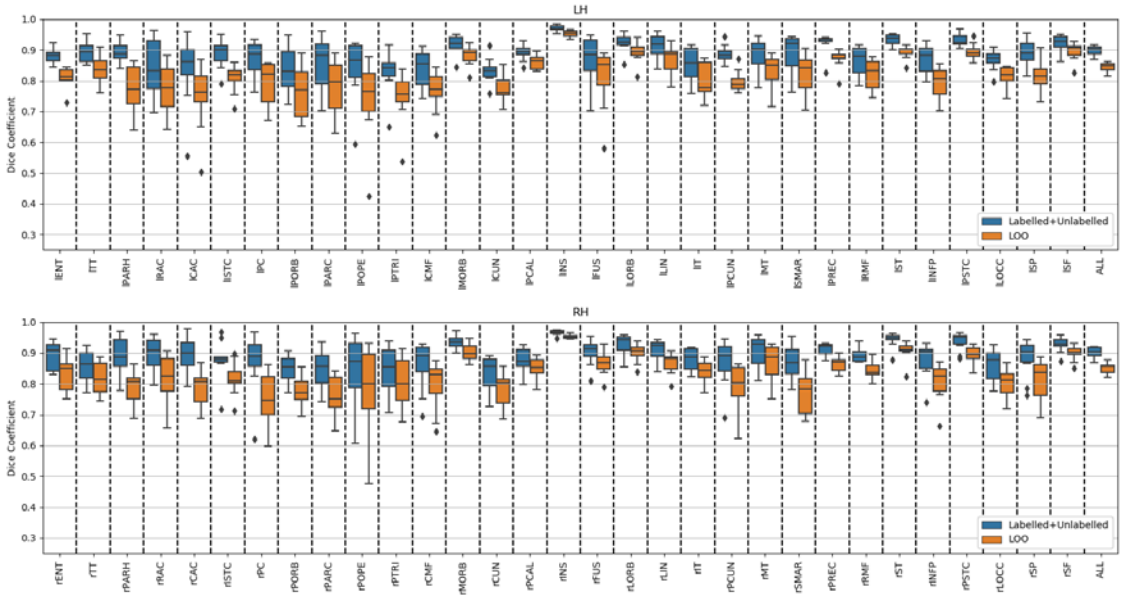


Figure 5: Dice coefficients comparing automatic labelling to manual parcellations of the M-CRIB(DKT) labels for labelled and unlabelled (L+U) and leave one out (LOO) curvature templates. Regions are sorted by area in ascending order. Whole-hemisphere accuracy is shown in the rightmost column (ALL). r= right, l = left, ENT = entorhinal, TT = transverse temporal, PARH = parahippocampal, RAC = rostral anterior cingulate, CAC = caudal anterior cingulate, ISTC = isthmus cingulate, PC = posterior cingulate, PORB =pars orbitalis , PARC = paracentral, POPE = pars opercularis, PTRI = pars triangularis, CMF = caudal middle frontal, MORB = medial orbitofrontal, CUN = cuneus, PCAL = pericalcarine, INS = insula, FUS = fusiform, LORB = lateral orbitofrontal, LIN = lingual, IT = inferior temporal, PCUN = precuneus, MT = middle temporal, SMAR = supramarginal, PREC = precentral,

RMF = rostral middle frontal, ST = superior temporal, INFP = inferior parietal, PSTC = postcentral, LOCC = lateral occipital, SP = superior parietal, SF = superior frontal.

Figure 6 displays label overlap scores for the M-CRIB-S(DK) parcellation scheme as well as the overlap measures for the automatically projected UNC DK labels compared with the manual M-CRIB labels. Whole-hemisphere accuracies were 0.89 (left) and 0.90 (right) for L+U, 0.84 (both left and right) for LOO for the M-CRIB-S(DK) atlas and 0.76 for both left and right for the UNC atlas. Generally, the M-CRIB-S(DK) accuracy rates were higher than for the UNC-based parcellations with mean regional scores exceeding UNC by 0.19 (LH) and 0.20 (RH) for L+U and by 0.12 (LH) and 0.14 (RH) for LOO. Average regional Dice coefficients were 0.87 (L+U), 0.80 (LOO) for the M-CRIB-S(DK) labels and 0.67 for the UNC atlas.

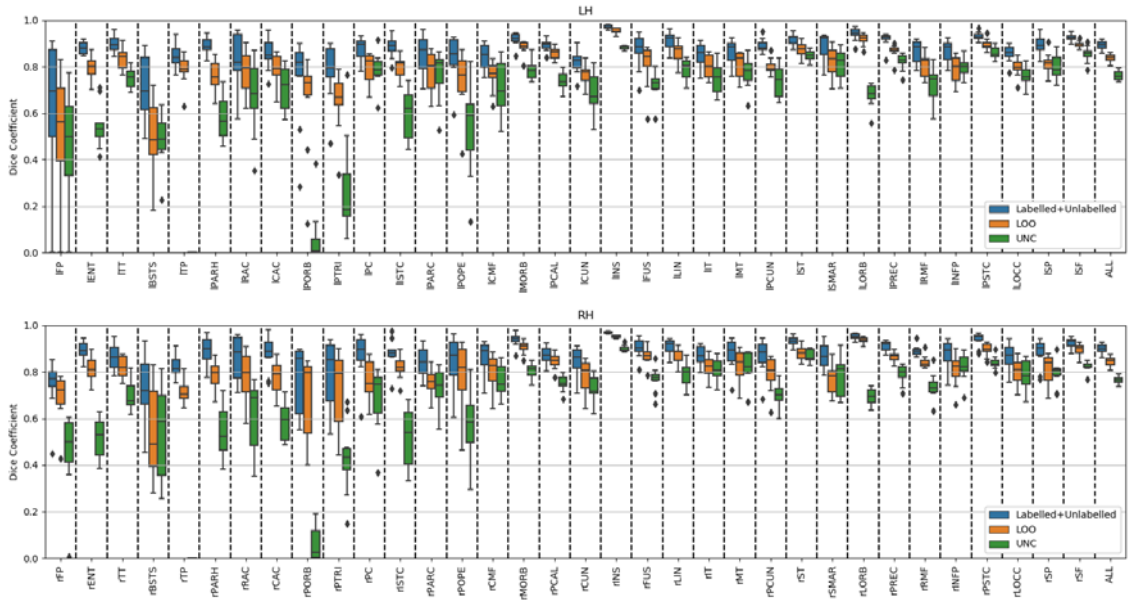


Figure 6: Dice coefficients comparing automatic labelling to manual parcellations of the M-CRIB(DK) labels for L+U and LOO curvature templates and the UNC atlas. Regions are sorted by area in ascending order. Whole-hemisphere accuracy is shown in the rightmost column (ALL). r= right, l = left, ENT = entorhinal, TT = transverse temporal, PARH = parahippocampal, RAC = rostral anterior

cingulate, CAC = caudal anterior cingulate, ISTC = isthmus cingulate, PC = posterior cingulate, PORB =pars orbitalis , PARC = paracentral, POPE = pars opercularis, PTRI = pars triangularis, CMF = caudal middle frontal, MORB = medial orbitofrontal, CUN = cuneus, PCAL = pericalcarine, INS = insula, FUS = fusiform, LORB = lateral orbitofrontal, LIN = lingual, IT = inferior temporal, PCUN = precuneus, MT = middle temporal, SMAR = supramarginal, PREC = precentral, RMF = rostral middle frontal, ST = superior temporal, INFP = inferior parietal, PSTC = postcentral, LOCC = lateral occipital, SP = superior parietal, SF = superior frontal.

Discussion

The primary contribution of this work is the provision of atlases and tools that facilitate cortical surface extraction and parcellation of the cortex into 31 to 34 regions based on T_2-weighted images of neonates around term equivalent age using a commonly used adult-compatible parcellation scheme, but with neonatal-derived training data. This will provide accurate age-specific cortical measures for neonates that out-perform those available from neonatal cortical atlases derived from adult training data. This work extends our previous M-CRIB and M-CRIB 2.0 volume-based atlases to enable cortical surface-based parcellation.

The age range of subjects suitable for processing under the proposed pipeline depends on whether tissue intensity contrast is adequate to reliably segment brain structures and extract cortical surfaces, and whether cortical folding complexity is enough to identify all macrostructural morphological features for region identification. The preprocessing pipeline involving tissue segmentation and cortical surface extraction using *DrawEM* and *Deformable* is tuned to the neonatal T_2-weighted white/grey matter contrast visible between 1 - 5 months (Wang et al., 2012) and has been demonstrated on data acquired between 34.4 - 43 weeks PMA (Bozek et al., 2018). Small-scale cortical folding occurs mostly late in gestation, with many sulci and gyri that define areal boundaries within the M-CRIB-S(DK) and M-CRIB-S(DKT) atlases not being reliably detectable until approximately

40 weeks PMA (Shi et al., 2011). It is possible that a sulcus, for example, may not have formed a sufficiently deep groove to be separated from neighbouring gyri. Thus, careful inspection of parcellation results for subjects scanned at ages below 40 weeks' PMA is required to ensure that the morphological features that define areal boundaries are present. From 5 to 8 months' PMA the T_2-weighted grey/white matter contrast is transitioning to become like that of an adult by about 9 months' PMA via an isocontrast phase (6 - 8 months) (Wang et al., 2012). The preprocessing pipeline is tuned to T_2-weighted contrasts and thus tissue segmentation would only be expected to work optimally up to approximately 5 months of age, before the T_2-weighted image becomes isointense. However, the parcellation component may be appropriate for subjects beyond 5 months of age, provided cortical surface data were obtained via other preprocessing means.

Parcellation accuracy of the automated surface-based labelling methods using the M-CRIB-S atlases was quantified according to overlap with manual labelling using Dice Coefficients for both the M-CRIB and M-CRIB 2.0 parcellation schemes. There were generally higher Dice Coefficients for larger regions. Quantitatively, there is an expected positive correlation between the size of a region and overlap score. According to Figure 4, the errors were largely concentrated close to region boundaries. This indicates only a small discrepancy in the location of the label boundaries, and we therefore considered our label placement to be reasonably accurate. There were few errors concentrated around the centre of labels, and if such errors were visible this would indicate inaccurate label placement due systematic or underlying differences in the choice of region definition, or registration problems. Within the M-CRIB-S(DK) atlas, regions that performed particularly poorly with lower overlap scores than for regions of similar size included the FP, BSTS and TP regions. Interestingly, the DKT protocol updated the DK scheme by removing the FP, BSTS and TP regions due to unclear boundary definitions, and low automatic parcellation reliability (Klein and Tourville, 2012). Thus, our results would support this decision. Globally, the LOO scores are slightly lower than the L+U scores, likely due to the LOO curvature templates being derived from

fewer subjects, and the test subject being left out of the training set. The original DKT atlas publication provided a Dice Coefficient-based overlap analysis based on the 42-subject subset used for that atlas, with a mean accuracy 0.916 (range 0.74 –0.99). The M-CRIB-S(DKT) atlas achieved comparable results of mean accuracy of 0.83 (range 0.83 – 0.97) for L+U analyses and mean 0.83 (range 0.74 – 0.95) for the LOO method. The original DK atlas publication (Desikan et al., 2006) utilised a jackknife Leave-One-Out method to compute Intraclass Correlation Coefficients on regional areas only to assess accuracy. Measures of overlap were not provided and thus cannot be compared with the M-CRIB-S(DK) atlas.

The relatively low Dice scores for the UNC parcellation shown in Figure 6, and the high mismatch in some regions seen for this parcellation in Figure 4, may be attributable to multiple factors that are, to some extent, non-delineable. Firstly, this may partly reflect the use of a single parcellated atlas image (as is employed in the UNC parcellation), compared with the use of a multiple-subject training set employed in the M-CRIB-S parcellations. The use of a training set confers the ability to use Bayesian methods to incorporate prior models in the training set with likelihood models generated from data to achieve a maximum *a posteriori* labelling. Secondly, a contributing factor to the observed mismatch may theoretically be an imprecise equivalence in interpretation of the original DK label descriptors between the original adult DK atlas, and the neonatal M-CRIB atlas. Indeed, the development of the M-CRIB atlases included specification of additional boundary descriptors. Additionally, the M-CRIB regions were traced volumetrically, whereas the adult DK regions were defined in a combination of volume and surface space. Due to the differences in morphology between neonatal and adult brains, it is not possible to test or quantify any mismatch arising from this potential factor, especially considering that some of the boundaries in the DK atlas were arbitrarily defined and may shift in unmeasurable ways during development. Relatedly, a third factor that may contribute to the areas of high mismatch between the UNC-labelled data and the manually labelled M-CRIB data is likely to be the use of an adult template projected into neonatal space, given

the changes in both gross and subtle brain morphology between these developmental time points. These morphology changes accompanying development fundamentally underlie the need for age-specific templates. Thus, despite the necessary challenges imposed by translation of boundaries, the current atlases constitute valuable datasets that are age-appropriate, precisely defined, and able to be applied probabilistically in surface-based pipelines.

The dHCP set of tools are currently available for cortical surface extraction and lobar parcellation. Makropoulos et al. recently highlighted the potential value of incorporating M-CRIB parcellation in these tools (Makropoulos et al., 2018), in that its compatibility with the DK (Desikan et al., 2006) adult parcellation may facilitate comparisons of cortical measures between the perinatal and adult time points. Here, we provide a publicly available, surfaced based cortical parcellation that can accomplish this objective.

The value of this surface-based atlas and the associated processing scripts is automated parcellation of the neonatal cortex that is straightforward to employ in longitudinal studies. The processing scripts and the M-CRIB-S(DK) and M-CRIB-S(DKT) atlases were constructed to be used with FreeSurfer, to produce compatible output and give a direct correspondence between region-based statistics such as cortical thickness, surface area, and curvature measures at neonatal, childhood and adult timepoints.

Conclusion

This paper presented the M-CRIB-S(DK) and M-CBRIB-S(DKT) atlases: surface-based versions of the volumetric M-CRIB and M-CRIB 2.0 atlases. It also presented an automated pipeline that involves segmentation of novel T_2-weighted neonatal images, extraction of cortical surfaces, followed by cortical parcellation with the M-CRIB-S(DK) and M-CRIB-S(DKT) atlases, which are neonatal versions of the adult DK and DKT atlases. The curvature template registration targets, average surfaces,

labelling training data, and pipeline execution scripts will be released. Additionally, for interoperability with the dHCP and UNC atlases we will also provide registered versions of the spherical template surfaces to be in correspondence to the dHCP and UNC 42-week curvature templates. Data will be available from our GitHub page (https://www.github.com/DevelopmentalImagingMCRI/M-CRIB-S).

Acknowledgements

We gratefully acknowledge support from members of the Victorian Infant Brain Studies (VIBeS) group, Developmental Imaging group, and Melbourne Children's MRI Centre at the Murdoch Children's Research Institute, and thank the families who participated in the study. This work was supported in part by the Australian National Health and Medical Research Council (NHMRC) (Project Grant ID 1028822 and 1024516; Centre of Clinical Research Excellence Grant ID 546519; Centre of Research Excellence Grant ID 1060733; Senior Research Fellowship ID 1081288 to P.J.A.; Early Career Fellowship ID 1053787 to J.L.Y.C., ID 1053767 to A.J.S., ID 1012236 to D.K.T.; Career Development Fellowship ID 1108714 to A.J.S., ID 1085754 to D.K.T.), Murdoch Children's Research Institute Clinical Sciences Theme Grant, the Royal Children's Hospital, the Department of Paediatrics at the University of Melbourne, the Victorian Government Operational Infrastructure Support Program, and The Royal Children's Hospital Foundation.

References

Alexander, B., Loh, W.Y., Matthews, L.G., Murray, A.L., Adamson, C., Beare, R., Chen, J., Kelly, C.E., Anderson, P.J., Doyle, L.W., Spittle, A.J., Cheong, J.L.Y., Seal, M.L., Thompson, D.K., 2018. Desikan-Killiany-Tourville Atlas Compatible Version of M-CRIB Neonatal Parcellated Whole Brain Atlas: The M-CRIB 2.0. BioRxiv.

Alexander, B., Murray, A.L., Loh, W.Y., Matthews, L.G., Adamson, C., Beare, R., Chen, J., Kelly, C.E., Rees, S., Warfield, S.K., Anderson, P.J., Doyle, L.W., Spittle, A.J., Cheong, J.L., Seal, M.L., Thompson,

D.K., 2017. A new neonatal cortical and subcortical brain atlas: the Melbourne Children's Regional Infant Brain (M-CRIB) atlas. NeuroImage 147, 841-851.

Amunts, K., Zilles, K., 2015. Architectonic Mapping of the Human Brain beyond Brodmann. Neuron 88, 1086-1107.

Beare, R.J., Chen, J., Kelly, C.E., Alexopoulos, D., Smyser, C.D., Rogers, C.E., Loh, W.Y., Matthews, L.G., Cheong, J.L., Spittle, A.J., Anderson, P.J., Doyle, L.W., Inder, T.E., Seal, M.L., Thompson, D.K., 2016. Neonatal Brain Tissue Classification with Morphological Adaptation and Unified Segmentation. Front Neuroinform 10, 12.

Bozek, J., Makropoulos, A., Schuh, A., Fitzgibbon, S., Wright, R., Glasser, M.F., Coalson, T.S., O'Muircheartaigh, J., Hutter, J., Price, A.N., Cordero-Grande, L., Teixeira, R., Hughes, E., Tusor, N., Baruteau, K.P., Rutherford, M.A., Edwards, A.D., Hajnal, J.V., Smith, S.M., Rueckert, D., Jenkinson, M., Robinson, E.C., 2018. Construction of a neonatal cortical surface atlas using Multimodal Surface Matching in the Developing Human Connectome Project. NeuroImage 179, 11-29.

Brett, M., Johnsrude, I.S., Owen, A.M., 2002. The problem of functional localization in the human brain. Nat Rev Neurosci 3, 243-249.

Coalson, T.S., Van Essen, D.C., Glasser, M.F., 2018. The impact of traditional neuroimaging methods on the spatial localization of cortical areas. Proc Natl Acad Sci U S A 115, E6356-E6365.

Desikan, R.S., Segonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968-980.

Fischl, B., 2012. FreeSurfer. NeuroImage 62, 774-781.

Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum Brain Mapp 8, 272-284.

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

Ghosh, S.S., Kakunoori, S., Augustinack, J., Nieto-Castanon, A., Kovelman, I., Gaab, N., Christodoulou, J.A., Triantafyllou, C., Gabrieli, J.D., Fischl, B., 2010. Evaluating the validity of volume-based and surface-based brain image registration for developmental cognitive neuroscience studies in children 4 to 11 years of age. NeuroImage 53, 85-93.

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., Smith, S.M., Van Essen, D.C., 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171-178.

Klein, A., Tourville, J., 2012. 101 labeled brain images and a consistent human cortical labeling protocol. Front Neurosci 6, 171.

Li, G., Wang, L., Shi, F., Gilmore, J.H., Lin, W., Shen, D., 2015. Construction of 4D high-definition cortical surface atlases of infants: Methods and applications. Med Image Anal 25, 22-36.

Loh, W.Y., Connelly, A., Cheong, J.L., Spittle, A.J., Chen, J., Adamson, C., Ahmadzai, Z.M., Fam, L.G., Rees, S., Lee, K.J., Doyle, L.W., Anderson, P.J., Thompson, D.K., 2016. A New MRI-Based Pediatric Subcortical Segmentation Technique (PSST). Neuroinformatics 14, 69-81.

Makropoulos, A., Aljabar, P., Wright, R., Huning, B., Merchant, N., Arichi, T., Tusor, N., Hajnal, J.V., Edwards, A.D., Counsell, S.J., Rueckert, D., 2016. Regional growth and atlasing of the developing human brain. NeuroImage 125, 456-478.

Makropoulos, A., Gousias, I.S., Ledig, C., Aljabar, P., Serag, A., Hajnal, J.V., Edwards, A.D., Counsell, S.J., Rueckert, D., 2014. Automatic whole brain MRI segmentation of the developing neonatal brain. IEEE Trans Med Imaging 33, 1818-1831.

Makropoulos, A., Robinson, E.C., Schuh, A., Wright, R., Fitzgibbon, S., Bozek, J., Counsell, S.J., Steinweg, J., Vecchiato, K., Passerat-Palmbach, J., Lenz, G., Mortari, F., Tenev, T., Duff, E.P., Bastiani, M., Cordero-Grande, L., Hughes, E., Tusor, N., Tournier, J.D., Hutter, J., Price, A.N., Teixeira, R.,

Murgasova, M., Victor, S., Kelly, C., Rutherford, M.A., Smith, S.M., Edwards, A.D., Hajnal, J.V., Jenkinson, M., Rueckert, D., 2018. The developing human connectome project: A minimal processing pipeline for neonatal cortical surface reconstruction. NeuroImage 173, 88-112.

Richards, J.E., Sanchez, C., Phillips-Meek, M., Xie, W., 2016. A database of age-appropriate average MRI templates. NeuroImage 124, 1254-1259.

Richards, J.E., Xie, W., 2015. Brains for all the ages: Structural neurodevelopment in infants and children from a life-span perspective. In: Benson, J.B. (Ed.), Advances in child developent and behavior. JAI, pp. 1-52.

Schuh, A., Makropoulos, A., Wright, R., Robinson, E.C., Tusor, N., Steinweg, J., Hughes, E., Grande, L.C., Price, A., Hutter, J., Hajnal, J.V., Rueckert, D., 2017. A deformable model for the reconstruction of the neonatal cortex. IEEE International Symposium on Biomedical Imaging, Melbourne, Australia, pp. p 800-803.

Shi, F., Yap, P.T., Wu, G., Jia, H., Gilmore, J.H., Lin, W., Shen, D., 2011. Infant brain atlases from neonates to 1- and 2-year-olds. PLoS One 6, e18746.

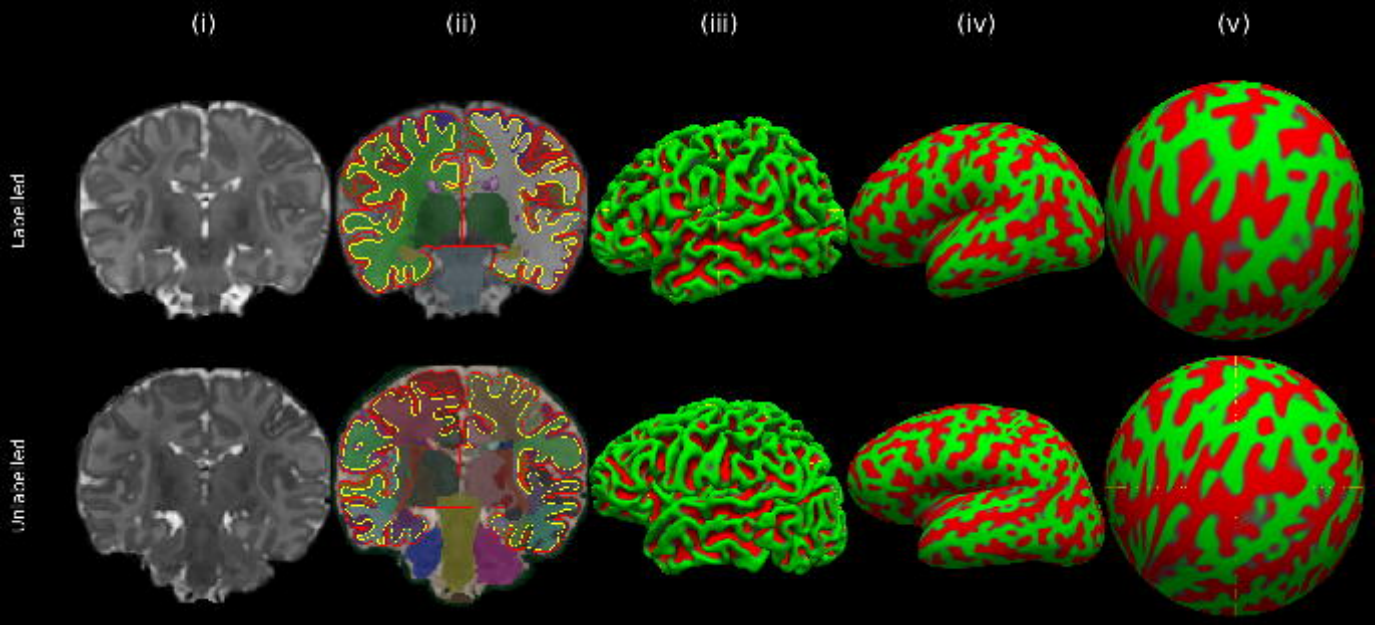
Spittle, A.J., Thompson, D.K., Brown, N.C., Treyvaud, K., Cheong, J.L., Lee, K.J., Pace, C.C., Olsen, J., Allinson, L.G., Morgan, A.T., Seal, M., Eeles, A., Judd, F., Doyle, L.W., Anderson, P.J., 2014. Neurobehaviour between birth and 40 weeks' gestation in infants born <30 weeks' gestation and parental psychological wellbeing: predictors of brain development and child outcomes. BMC Pediatr 14, 111.

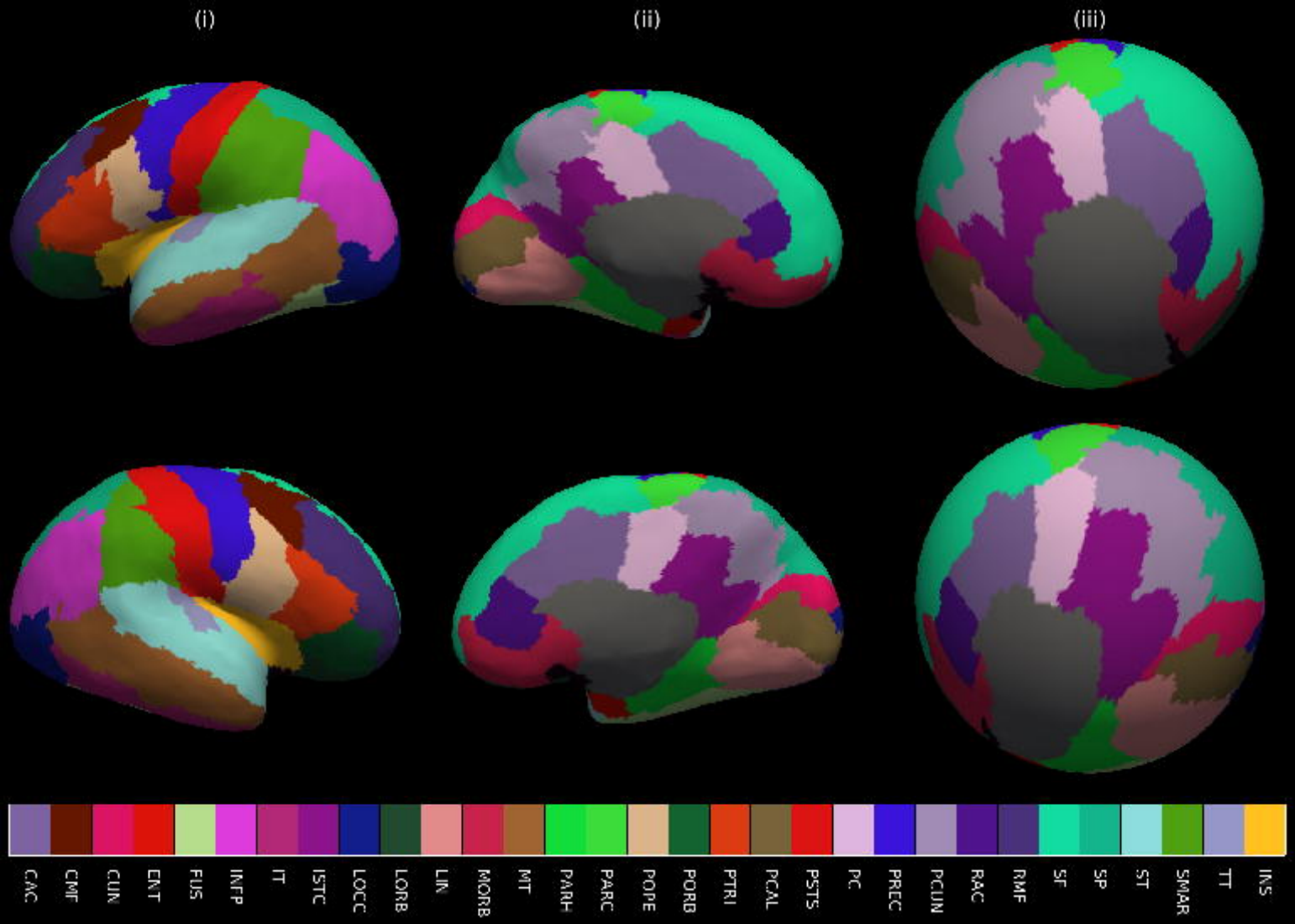
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15, 273-289.

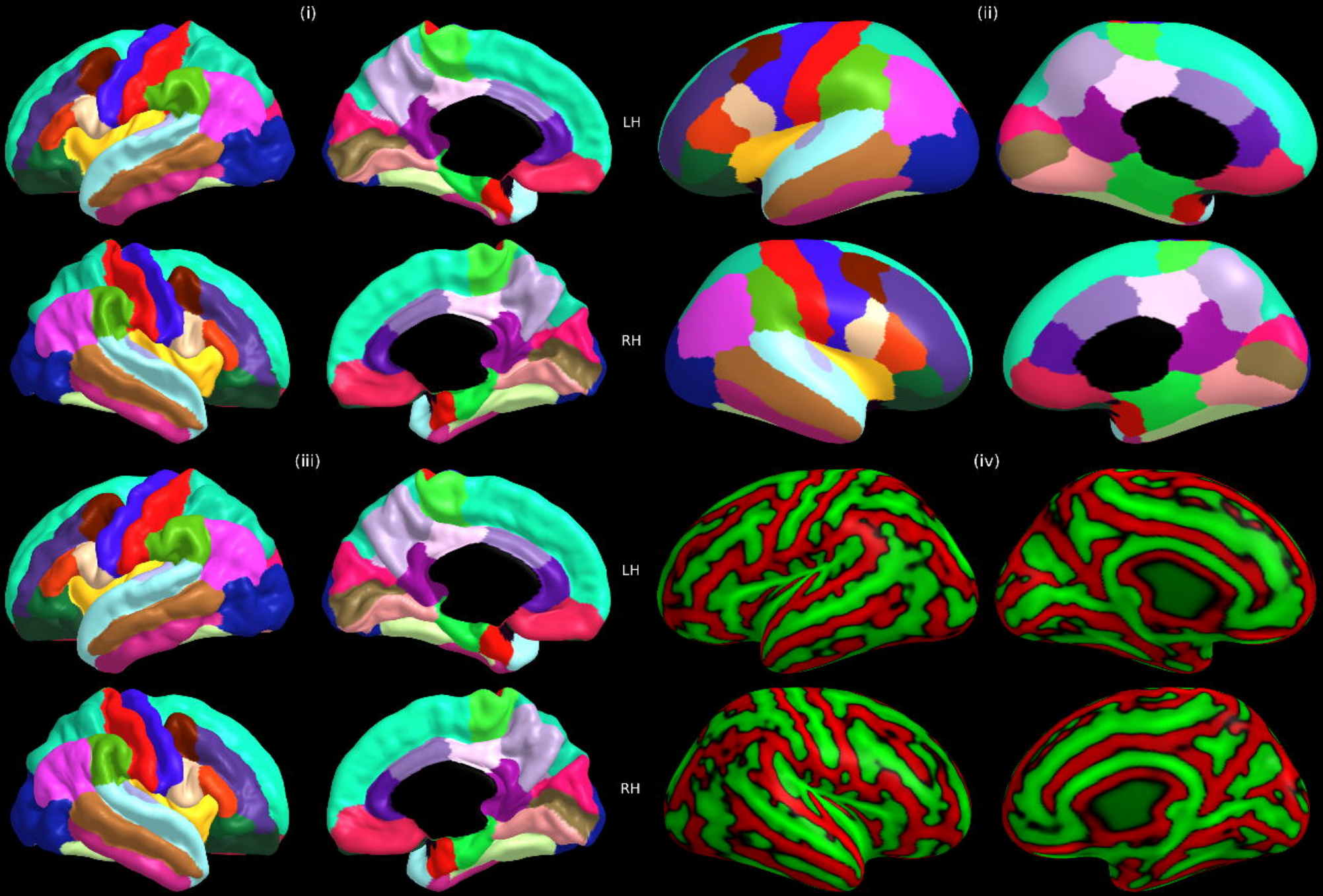
Walsh, J.M., Doyle, L.W., Anderson, P.J., Lee, K.J., Cheong, J.L., 2014. Moderate and late preterm birth: effect on brain size and maturation at term-equivalent age. Radiology 273, 232-240.

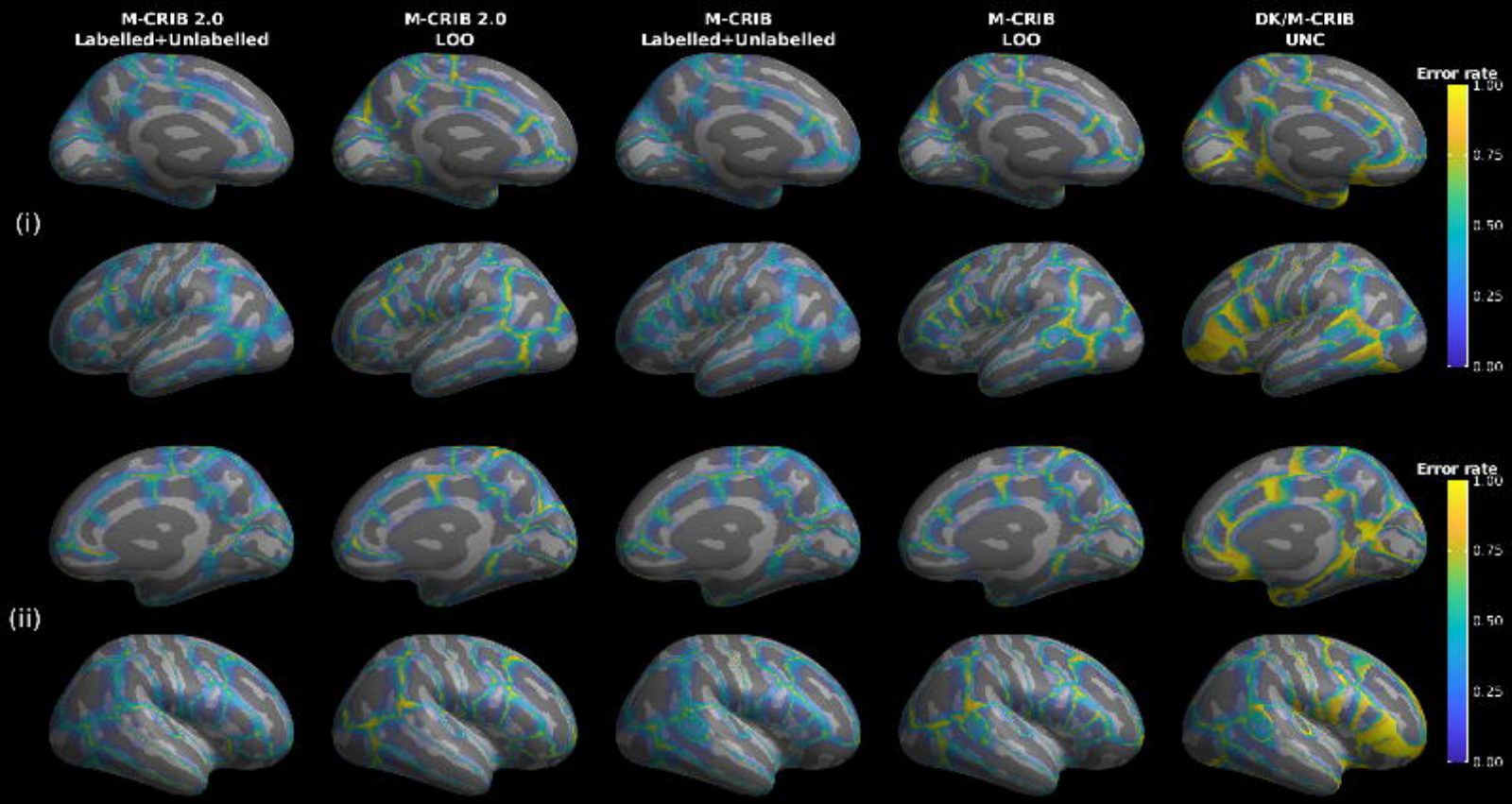
Wang, L., Shi, F., Yap, P.T., Gilmore, J.H., Lin, W., Shen, D., 2012. 4D multi-modality tissue segmentation of serial infant images. PLoS One 7, e44596.

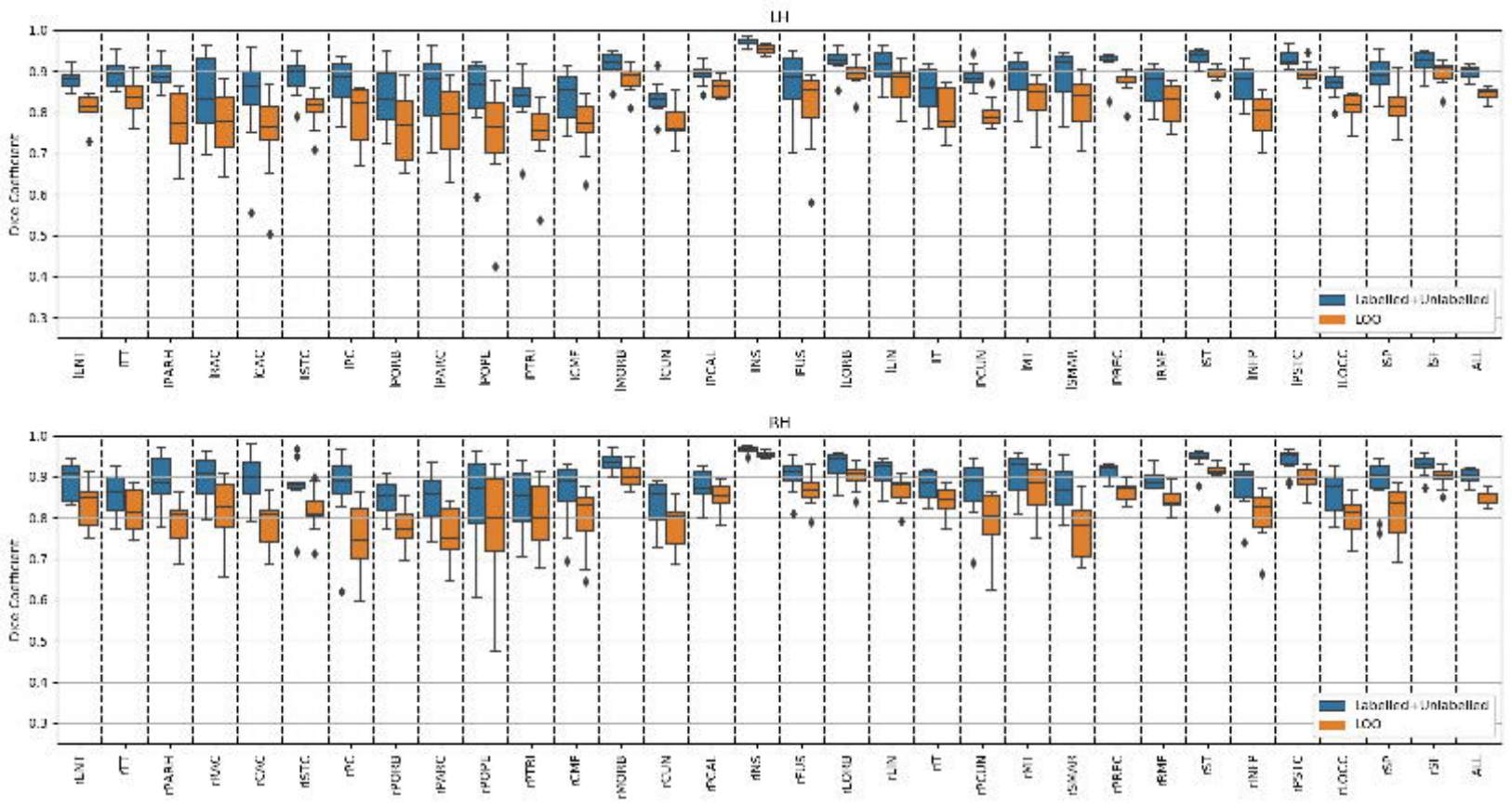
Wu, Z., Li, G., Wang, L., Lin, W., Gilmore, J.H., Shen, D., 2018. Construction of Spatiotemporal Neonatal Cortical Surface Atlases Using a Large-Scale Dataset. Proc IEEE Int Symp Biomed Imaging 2018, 1056-1059.

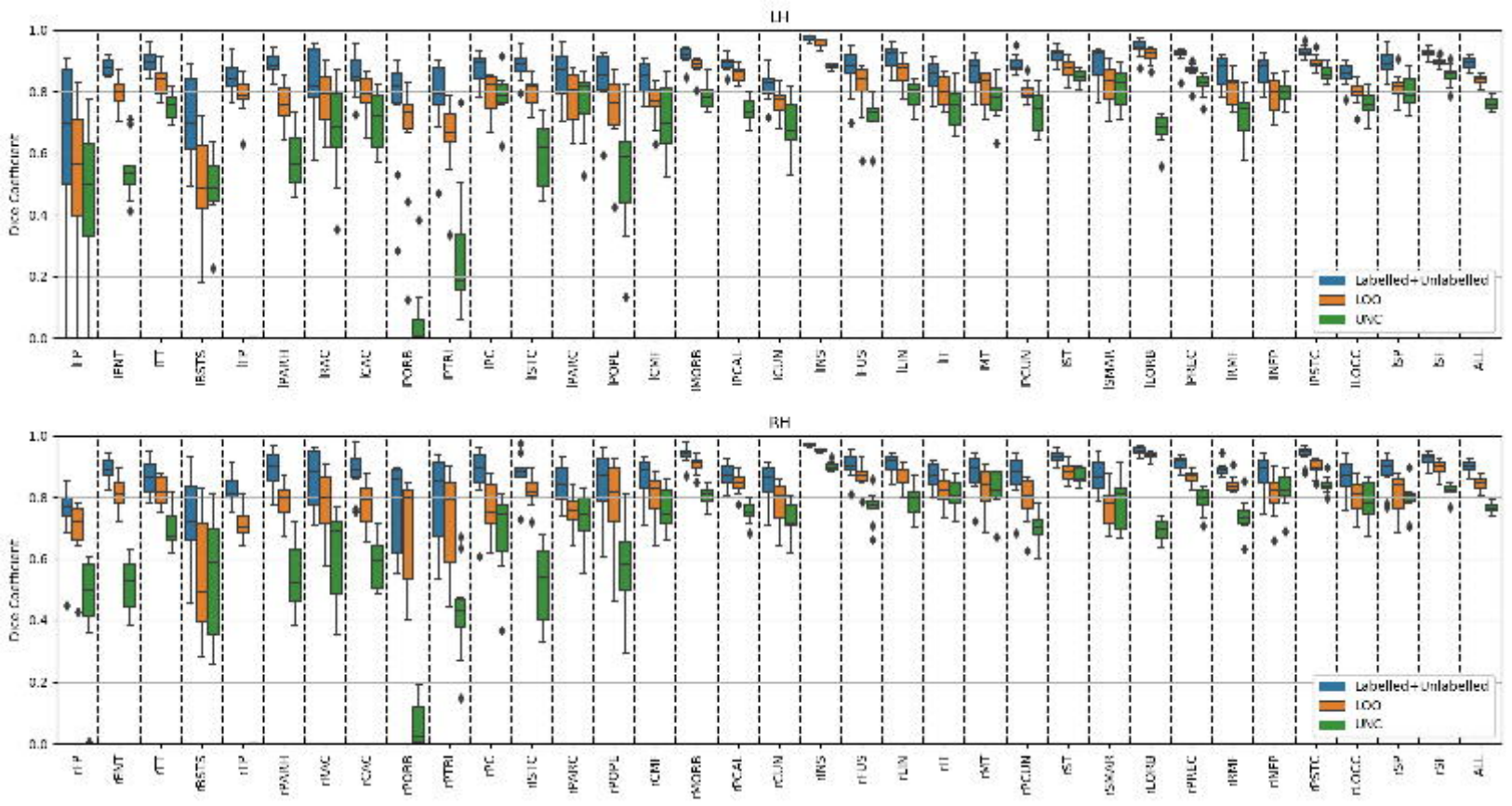

(i)
(ii)
(iii)
(iv)
(v)
Labelled
Unlabelled


(i)
(ii)
(iii)
CAC
CMF
CUN
ENT
FUS
INFP
IT
ISTC
LOCC
LORB
LIN
MORB
MT
PARH
PARC
POPE
PORB
PTRI
PCAL
PSTS
PC
PREC
PCUN
RAC
RMF
SF
SP
ST
SMAR
TT
INS


(i)
(ii)
LH
RH
(iii)
(iv)
LH
RH


M-CRIB 2.0
Labelled+Unlabelled
M-CRIB 2.0
LOO
M-CRIB
Labelled+Unlabelled
M-CRIB
LOO
DK/M-CRIB
UNC
Error rate
1.00
0.75
0.50
0.25
0.00
(i)
(ii)


LH
RH
Dice Coefficient
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Labelled + Unlabelled
LOO
lLVT
lTT
lPARH
lMOC
lCMC
lSTC
lPC
lPORB
lPARC
lPOPL
lPTRI
lCMF
lMORB
lCUN
lPCAL
lINS
lFUS
lLORB
lLIN
lIT
lPCUN
lMT
lSMAR
lPREC
lRMF
lST
lMTP
lPSTC
lLOCC
lSP
lSI
ALL
rLVT
rTT
rPARH
rMOC
rCMC
rSTC
rPC
rPORB
rPARC
rPOPL
rPTRI
rCMF
rMORB
rCUN
rPCAL
rINS
rFUS
rLORB
rLIN
rIT
rPCUN
rMT
rSMAR
rPREC
rRMF
rST
rMTP
rPSTC
rLOCC
rSP
rSI
ALL


LH
RH
Dice Coefficient
Labelled+Unlabelled
LOO
UNC
lFP
lFNT
lTT
lBSTS
lTP
lPAIUI
lRAC
lCAC
lPORB
lPTRI
lPC
lISTC
lPARC
lPOPL
lCMI
lMORB
lPCAL
lCUN
lINS
lFUS
lLIN
lIT
lMT
lPCUN
lST
lSMAR
lLORB
lPRLC
lRMI
lNFP
lPSTC
lLOCC
lSP
lSI
ALL
rFP
rFNT
rTT
rBSTS
rTP
rPAIUI
rRAC
rCAC
rPORB
rPTRI
rPC
rISTC
rPARC
rPOPL
rCMI
rMORB
rPCAL
rCUN
rINS
rFUS
rLIN
rIT
rMT
rPCUN
rST
rSMAR
rLORB
rPRLC
rRMI
rNFP
rPSTC
rLOCC
rSP
rSI
ALL