

INTRODUCTION

We very much appreciate the reviewers' constructive comments. To improve the rigor of our research strategy, we have updated the literature review, provided more preliminary data, improved the design of deep learning model implementation, increased the sample size, included multi-site data, and provided strategies for data harmonization. We reinforced biological hypotheses and clarified resource sharing plans for clinical application.

Significance/clinical application: Our preliminary results and other previous studies have demonstrated that cortical plate (CP) and subplate (SP) thickness can be regionally altered by defects in brain-expressed genes, developmental and genetic brain disorders, and hypoxic-ischemic brain injury. Regional analysis of CP and SP thickness infer regional alterations in cortical microstructure and in the development of cortico-cortical connections, providing mechanistic insight when correlated with neurodevelopmental outcomes. However, it is difficult to perform such studies on large samples due to the lack of tools that can reliably and automatically measure regional CP and SP thickness on *in vivo* fetal MRI. We propose to develop a new methodology and an **advanced resource sharing plan for clinical translation**. Our algorithms will be packaged in docker containers and distributed in our public Github repository. We will distribute a fully automated pipeline using **ChRIS** (Children's Research Integration System), a web-based framework that enables MRI data to be processed in a clinical setting. **Dr. Pienaar** (lead of the ChRIS project), will be responsible for our software deployment.

Biological hypotheses: To perform a more hypothesis-driven analysis, we now focus on two (instead of four) abnormal groups: **polymicrogyria (PMG)** and **agenesis of the corpus callosum (ACC)**. Hypotheses for CP and SP thickness alterations in PMG and ACC is supported by our preliminary data and prior studies. PMG has abnormal cortical layer structure with reduced neuronal populations and thickness of the cortex in histopathological studies. Cortical thickness alterations have also been found in adolescents with PMG on MRI. Since ACC lacks interhemispheric cortico-cortical connections, a decrease in SP thickness is hypothesized. In addition, our preliminary data showed disorganized sulcal folding patterns in fetuses and children with PMG and ACC¹⁻⁴. Since cortical thickness impacts cortical folding, disrupted sulcal patterns may be related to an altered spatial pattern of cortical thickness in PMG and ACC. To ensure accurate interpretation of our findings, we now include 10% effort of **Dr. Grant**, a highly qualified pediatric neuroradiologist.

Increased sample size: To make our methods more generalizable and fully characterize the temporal changes of CP and SP thickness in normal and abnormal groups, multi-site fetal MRI datasets will be used. We have increased the sample size of not only typically developing (TD) fetuses but also fetuses with cerebral abnormalities through collaboration with Children's Hospital Pittsburgh (Dr. Panigrahy, Co-I) and Tufts Medical Center (Dr. Tarui, Co-I). We have also updated our database of fetal MRI cases with ACC at Boston Children's Hospital (total *n*: 280 controls, 70 PMG, 50 ACC). For **data harmonization** between sites, we will first apply two intensity normalization methods (Z-score, piecewise linear histogram matching) and also use a recent technique, ComBat, to statistically harmonize CP and SP thickness from three different sites. ComBat has been used for the multi-site/scanner analyses of cross-sectional and longitudinal cortical thickness postnatally.

SP segmentation on T2-weighted *in vivo* MRI: We have expertise in the definition of the SP with fetal neuroanatomical knowledge on T2-weighted *in vivo* fetal MRI⁵⁻⁷. Our extensive experience in SP segmentation will enable us to construct high-quality ground truth data. Since a deep learning method utilizes spatial and other complex features as well as intensity features of brain structures, it can **automatically segment the SP in spite of subtle tissue contrast in MRI**. Using our SP segmentation images of 20 TD fetuses, we trained our deep learning architecture, previously developed for fetal CP segmentation, and attempted to automatically segment the SP. Our preliminary data show very promising results and support the feasibility of accurate fetal SP segmentation.

Deep learning model: We have clarified details of our deep learning model implementation and validation (network architecture, training strategy, model optimization, external cross validation). We will **implement and evaluate both 3D convolutional neural network (CNN) and 2D CNN with multi-view aggregation**. We propose a biologically inspired data augmentation using deformable image registration between age-specific fetal brain atlases. For validation, we will compare our augmentation with standard geometry transformation augmentation.

Imaging protocols: T2-weighted HASTE imaging is the primary method for clinical evaluation of the fetal brain. Our preliminary results show that **accurate CP and SP segmentation is feasible using HASTE images**. A recent study reported that EPI-FLAIR sequence could improve SP visualization in fetal MRI, but this is not routine at many centers (e.g. BCH, CHP, TMC). For alternative strategies and future work, we will perform FLAIR imaging for prospective fetal projects and test if they enhance the performance of the SP segmentation model.

Data quality control: We will use an **automated method for scoring the quality of raw fetal MR images in terms of motion artifacts and noise**, which has been recently developed using a deep CNN model. We will **classify low-score images, visually recheck them for exclusion, and then perform motion correction**. The motion-corrected volumes will also be inspected and excluded if motion artifacts and image noise still remain.

SPECIFIC AIMS

The cortical plate (CP) and subplate (SP) are transient fetal brain compartments where post-migratory cortical neurons differentiate and early synaptic interaction and guidance of axons occur⁸⁻¹⁰. CP thickness is thought to be related to the number and size of cells within a column, packing density, intracortical myelin, and synapses¹¹⁻¹⁵. SP thickness is associated with the amount of thalamic and cortical afferents and cortico-cortical connections^{8,16-19}. Thickness variations of the CP and SP have been identified across different regions and gestational ages by manual segmentation of MRI data and shown to be associated with regional cortical gene expression in postmortem *ex vivo* and *in vivo* fetal brains^{7,10,17}. CP layer structure and thickness are also altered by developmental brain disorders^{15,20,21} and SP neuronal death occurred due to perinatal hypoxic-ischemic brain injury¹⁹ in human histopathological and animal model studies, suggesting regional alterations in CP and SP thickness can be present prenatally. Hence, the identification of abnormal CP and SP thickness may be important for a better understanding of the mechanisms underlying altered brain development. However, *our knowledge and research of human in utero CP and SP thickness remain limited due to the lack of reliable tools that can automatically measure regional CP and SP thickness from in vivo fetal brain MRI.*

Automatic cortical gray matter segmentation and thickness estimation between inner and outer cortical surfaces have been widely used for analyzing postnatal cortical thickness with MRI²²⁻²⁹. However, compared to child or adult brains, fetal brains are much smaller in size and have different tissue contrast in MRI. Fetal brain MRI also shows lower effective resolution and suffers from head motion, which causes artifacts. This makes it difficult to segment thin CP and define the boundary between CP and cerebrospinal fluid (CSF) (outer CP boundary) in highly folded, deep, and narrow sulcal regions. Moreover, the contrast between SP and intermediate zone is low, and the surface of inner SP boundary has different shape complexity compared to CP surface. Thus, it is *challenging to extract accurate CP and SP surfaces and define geometrically appropriate thickness between the CP and SP surfaces. This project will overcome these barriers and develop a fully automatic pipeline to extract regional CP and SP thickness that includes deep learning and deformable surface model methods.* Next, we will *examine regional CP and SP thickness growth from 18 to 37 gestational weeks (GW) in typically developing (TD) fetuses.* Finally, since malformations of cortical development and disrupted cortico-cortical connections may alter CP and SP thickness, we will *explore changes in CP and SP thickness in fetuses with two different cerebral abnormalities, polymicrogyria (PMG) and agenesis of the corpus callosum (ACC).*

Aim 1. Develop deep learning methods for CP, SP, and sulcal CSF segmentation from fetal brain MRI

Generalized deep convolutional neural network (CNN) models will be developed using multi-site fetal MRI data for automatic CP and SP segmentations. CSF in deep sulcal regions will be detected using a patch-based CNN for accurate outer CP surface reconstruction. We aim to have the Dice coefficient greater than 90% and mean surface distance error less than 0.18mm between automatic and manual segmentations.

Aim 2. Develop methods for outer CP and inner SP surface reconstruction and measurement of regional CP and SP thickness in the fetal brain

Inner CP surface will first be reconstructed, and then outer CP and inner SP surfaces will be extracted using optimal surface deformation algorithms that can be generally applied to a wide range of gestational age. CP and SP thickness will be measured based on vertex-wise correspondence between all CP and SP surfaces. Reliability and sensitivity tests will be performed using different imaging subsets within the same subject and artificial data created by moving the CP and SP boundary.

Hypothesis a: No significant intra-subject differences in CP and SP thickness will be found between the subsets.

Hypothesis b: Automatic CP and SP thickness measure will be significantly correlated with the cortical volume ratio that varies when the CP and SP are artificially adjusted.

Aim 3. Define spatiotemporal patterns of CP and SP thickness in fetal brains from 18 to 37 GW

The temporal changes of regional CP and SP thickness will be modeled as a function of GW in TD fetuses and fetuses with PMG (cortical malformation) and ACC (disrupted cortico-cortical connectivity) and compared between the groups.

Hypothesis a: The growth rate of CP thickness and/or the GW of maximum growth rate will be regionally variable in TD fetuses and fetuses with PMG and ACC will show regional alterations compared to TD fetuses.

Hypothesis b: The growth rate of SP thickness and/or the maximum SP thickness will be regionally variable in TD fetuses and fetuses with PMG and ACC will show regional alterations compared to TD fetuses.

This study will lay the foundation for a novel biomarker that will provide greater insight into the mechanisms of normal and altered *in utero* brain development, allowing assessment of the impact of genetic defects and maternal exposures and informing strategies for future *in utero* therapies. Our methods that will be developed in the proposed study will be publicly distributed using a web-based neuroimage computation platform, which will facilitate translation of fetal CP and SP thickness analysis at other clinical research centers.

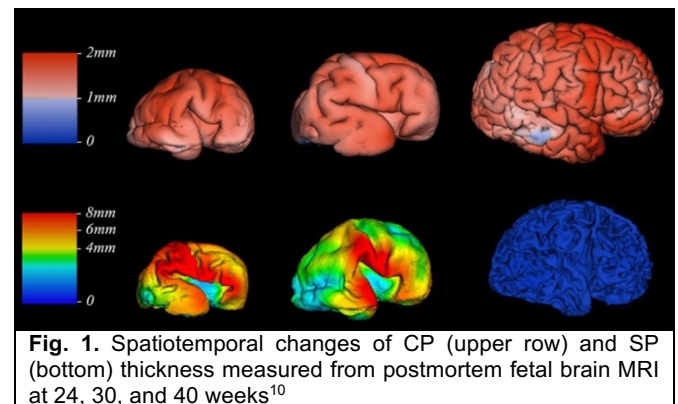
RESEARCH STRATEGY SIGNIFICANCE

The thickness of the human cerebral cortex has been extensively investigated using *in vivo* MRI. Cortical thickness postnatally estimated with MRI is variable across different regions and individuals and associated with age³⁰, sex³¹, and intellectual ability³² in healthy subjects as well as various neurodevelopmental, neurodegenerative, and psychiatric disorders^{11,12,33-41}. Postnatal cortical thickness measurement has contributed greatly to our understanding of cortical development across the lifespan and cognitive function and provided mechanistic insight into disease onset and progression in various brain disorders. As *variations in regional cortical thickness influenced by genetic and environmental factors and developmental disorders* have been found to be detectable in the early neonatal period (from 6-9 days)^{11,12,33,40,41}, these variations *likely originate in utero*. In the human fetal brain, cortical plate (CP) thickness is associated with neural developmental events such as the number and size of cells within a column, packing density, intracortical myelin, and synapses¹¹⁻¹⁵, and may be altered by defective brain development *in utero*. One of the most common cortical malformations, **polymicrogyria (PMG)**, is caused by abnormal prenatal post-migrational development, and shows abnormal cortical lamination as well as reduced neuronal populations and cortical thickness in histopathological studies^{21,42,43}. Other brain disorders also have altered laminar structures and cortical cell density associated with CP thickness changes on histopathology^{15,42}. Hence, human *in utero* CP thickness is likely altered in various developmental and genetic brain disorders, but it is largely unknown.

The subplate (SP) is a transient cytoarchitectonic compartment beneath the CP in the fetal cerebrum and composed of SP neurons and “waiting” thalamocortical and cortico-cortical fibers⁸. The volume of SP increases with age between 13 and 30 gestational weeks (GW) in most areas¹⁰. The *maximal thickness of SP* during this period may reflect an increased amount of cortical afferents within SP, which form transient synapses before continuing into the cortex, and may be *related to the amount of cortico-cortical connections*^{8,16,18,44}. The SP structure could be influenced by atypical brain development *in utero*. Brain malformations with a lack of cortico-cortical connections such as **agenesis of the corpus callosum (ACC)** may cause a decrease in SP thickness. SP neurons were also shown to be selectively vulnerable to early environmental exposures such as hypoxic-ischemic brain injury in animal models, which may result in defective thalamic and cortical afferents and alter SP thickness¹⁹. Moreover, the volumes of the CP and SP were positively correlated within the same areas¹⁰. Thus, *alterations in CP and SP thickness may occur during fetal periods, reflecting altered cortical microstructure and altered cortico-cortical connectivity development*^{8,16,45,46}.

Alterations in CP and SP thickness may occur in a region-specific manner.

Brain developmental processes (neuronal proliferation, migration, and differentiation, and the growth of dendrites and synapses) are strongly associated with dynamic spatiotemporal patterns of gene expression. The majority of brain-expressed genes show large variabilities in expression between different cortical regions and transient compartments (ventricular, subventricular, intermediate, SP, and CP zones) in the human fetal brain⁴⁷⁻⁴⁹. We have revealed spatial variability of CP and SP thickness (Fig. 1) and relationships between mean regional SP/CP thickness ratio and expression level of genes which play an essential role in neuronal migration^{7,10}. Thus, CP and SP structural alterations influenced by defects in brain-expressed genes may show a region-specific pattern. Human brain specimens who experienced environmental exposures such as a reduction of cerebral oxygen supply during gestation showed impairment of cortical growth and connectivity in specific sub-regions of the frontal cortex⁵⁰. *Regional analysis of CP and SP thickness could be used to infer regional differences in neuronal structure and connections and their maturation that relate to mechanisms of alteration.*



The human CP and SP thickness has been measured by histological and MRI studies of the postmortem brain, which do not translate to living patient care. Although there are *in vivo* MRI studies to measure CP and SP thickness *in utero*^{7,17}, the CP and SP zones are defined by manual segmentation, which is time-consuming, subject to intra- and inter-rater variabilities, and not feasible for large clinical studies. Thus, *automatic measures of regional CP and SP thickness in utero from in vivo fetal MRI are needed.* *The definition of spatiotemporal patterns of CP and SP thickness using a large dataset of fetal brain MRIs across a wide gestational age range* could provide a baseline for comparison to abnormal fetal brain development. It will lay the foundation for a *novel*

biomarker that can lead to greater insight into the mechanisms underlying normal and altered in utero brain development. It has potential not only for early detection of abnormal brain development but may also enable detection of the impact of genetic and environmental factors on early brain development, informing strategies for type and timing of future *in utero* therapies.

Well-known image processing pipelines such as CIVET^{22,23,25} and Freesurfer^{24,51,52} automatically measure cortical thickness using the white and gray matter surfaces from T1-weighted structural images for child or adult brains. However, fetal brains are much smaller in size and have different transient cellular compartments and different tissue contrast in MRI. Unlike adult brains, ***single shot spin-echo T2-weighted imaging is the primary method*** for clinical evaluation of fetal brain structure as it minimizes fetal motion and provides relatively good tissue contrast and detail of fetal brain anatomy⁵³. Fetal brain compartments including CP and SP regions have been identified and analyzed in several *in vivo* T2-weighted MRI studies^{5-7,17,54-56}. Furthermore, fetal brain MRI shows lower effective resolution and suffers from motion artifacts. Thus, ***previous traditional tools cannot be utilized to automatically process fetal brain MRIs and measure CP and SP thickness.***

INNOVATION

The proposed research will be ***the first study to develop a pipeline for measuring and evaluating regional CP and SP thickness in the human fetal brain using in vivo MRI that is applicable across a wide gestational age range.*** Spatiotemporal patterns of fetal CP and SP thickness will be explored from the second trimester onwards. Given that fetal brain MRIs are less common than pediatric and adult brain MRIs and there is no public dataset of fetal brain MRIs, *the size of our multi-site fetal MRI datasets is relatively large (total n=400).*

We will provide novel techniques to process fetal MRI and calculate CP and SP thickness. Our pipeline will include ***deep learning-based CP and SP segmentation with identification of sulcal cerebrospinal fluid (CSF) regions; accurate CP and SP surface extraction and geometrically appropriate thickness measurement;*** and surface-based local thickness analysis. It is crucial to extract the deep and narrow sulcal CSF regions for accurate outer CP surface extraction and thickness estimation. However, it is more challenging in fetal brains compared to adult brains because the fetal brain is small and effective image resolution is relatively low. Also, the boundary between SP and the intermediate zone (inner SP boundary) has low contrast in MRI. *No studies have yet provided a method for automatic sulcal CSF and SP segmentation in the fetal brain.* Given that deep learning techniques have emerged as one of the most prominent approaches for image segmentation⁵⁷⁻⁵⁹, a state-of-the-art convolutional neural network (CNN) model will be designed as follows: (1) Our deep CNN model will adopt a ***novel hybrid loss function and multi-view aggregation approach with a test-time augmentation*** to improve segmentation accuracy⁶⁰. (2) ***For the estimation of sulcal CSF regions, we will design and use a patch-based CNN architecture.*** (3) To build a more robust and generalized segmentation tool that can be applied to other datasets, we will train our deep CNN using a relatively *large dataset of fetal brain MRIs collected from three different sites* with different scanners and field strengths and utilize a novel *biologically inspired data augmentation* approach. Based on accurate CP, SP, and sulcal CSF segmentation, we will develop a *deformable model method* to extract the outer CP and inner SP surfaces while *preserving correct surface topology and vertex-wise correspondence between all CP and SP surfaces*. Our surface extraction method will be *optimized and specialized for fetal brains and generalized by performing systematic evaluations using a large dataset.*

To observe and analyze the structural image of fetal brains, several sets of images in three orthogonal planes (axial, sagittal, and coronal views) are usually acquired and combined using image processing such as *slice-to-volume registration for motion correction and super-resolution volume reconstruction*⁶¹⁻⁶⁶. These imaging and additional processing steps may affect the sensitivity and reliability of fetal brain measures. Thus, we will *test the reliability and sensitivity of fetal CP and SP thickness measures.*

Development and validation of our automatic methods would allow us to process a large dataset with a minimal amount of manual work, which in turn will enable us to understand trajectories (temporal dynamics) and inter-subject variabilities of CP and SP thickness as well as reveal region-specific patterns of their growth. We will generate *the first model of typical in utero human CP and SP thickness development from early fetal life (second trimester) to near term age (18-37 GW).* We will also *explore alterations in CP and SP thickness growth in fetuses with neural developmental abnormalities:* PMG (abnormal post-migrational development and cortical malformation) and ACC (disrupted cortico-cortical connections linking the two hemispheres). It may reveal new findings and generate new hypotheses about abnormal cortical structural development in brain malformations during the fetal period.

We will distribute a fully automated pipeline for fetal CP and SP thickness measurement using our advanced

web-based medical data and computation platform called **ChRIS** (*Children's Research Integration System*)⁶⁷. ChRIS utilizes cloud technologies and handles the reception, storage, management, processing, and visualization of MRI data. It will allow us to perform clinical translational research as it makes processing of fetal MRI and assessment of *in utero* CP and SP thickness feasible in a clinical setting.

The feasibility of the proposed project is supported by our extensive experience in fetal MRI processing and surface-based cortical analysis^{1,3,31,68-71} and an expert team comprising of a bioengineer, fetal neuroradiologist, neuroanatomist, neurologist, biostatistician, software architect, and a collaborator who was a pioneer in the development of an image processing pipelines for cortical thickness analysis in pediatric and adult brains^{23,25,28,72}.

APPROACH

Retrospective cohort and MRI data

We propose a new methodology and analysis with a large dataset of retrospective fetal MRIs that were previously collected at Boston Children's Hospital (BCH), Children's Hospital of Pittsburgh (CHP), and Tufts Medical Center (TMC) (**280** typically developing [TD] fetuses and **120** fetuses with cerebral abnormalities). Fetal subjects were scanned at one time point. TD fetal MRIs were acquired by recruitment at BCH, and also collected retrospectively from clinical fetal MRIs performed to screen for abnormalities at BCH, CHP, and TMC but found to be normal (n [BCH/CHP/TMC]: 150/100/30). Inclusion criteria for TD fetuses included no serious maternal medical conditions (nicotine or drug dependence, morbid obesity, cancer, diabetes, and gestational diabetes), maternal age between 18-45 years, fetal gestational age between 18-37 weeks. Exclusion criteria included multiple gestation pregnancies, dysmorphic features on ultrasound (US) examination, brain malformations, or brain lesions on US, other identified organ anomalies on US, known chromosomal abnormalities, known congenital infections and any abnormality on the fetal MRI. Fetal MRIs that were performed between 18-37 GW for clinical neurological concerns were collected at BCH, CHP, and TMC. Based on our initial database, we will include **70 fetuses with PMG** (n [BCH/CHP/TMC]: 50/20/0) and **50 fetuses with ACC** (25/15/10). For all healthy and abnormal fetuses, relevant neuroimaging and clinical data will be reviewed, and their clinical/imaging diagnoses and demographic data will be reconfirmed by a fetal neuroradiologist, neurologist, and neuroanatomist. Pregnant women were scanned without maternal sedation using single shot spin-echo T2-weighted imaging⁷³ due to its capability to obtain single slice data quickly and minimize the effect of motion artifacts. After localizing the fetal brain, a total of 6-18 scans were acquired multiple times in three orthogonal orientations.

BCH: Siemens 3T, Half-Fourier Acquisition Single-Shot Turbo Spin-Echo (HASTE), 1mm in-plane resolution, 256×256 acquisition matrix, TR=1.5s, TE=120ms, slice thickness=2-4mm

CHP: Siemens 3T, HASTE, 1.1-1.4mm in-plane resolution, 256×256 acquisition matrix, TR=1.1s, TE=78ms, slice thickness=3mm

TMC: Phillips 1.5T, Multiplanar Single-Shot Turbo Spin-Echo, 1mm in-plane resolution, 256×256 acquisition matrix, TR=12.5s, TE=180ms, slice thickness=2-4mm

Automatic fetal MRI pre-processing

From raw fetal MRI data, only the **brain region** will be **extracted to exclude maternal and nonbrain tissues** using a **deep learning-based automatic segmentation tool** that we recently published⁷⁴. We trained a **2D U-Net architecture**⁷⁵ using ~15,700 slices from 397 multi-view triplane MRIs of 81 TD fetuses (28.0±4.3 GW) and obtained ~95% of accuracy **for automatic brain extraction** (Fig. 2)⁷⁴. The N4 bias field correction will be performed to correct for intensity inhomogeneity⁷⁶. We will filter the images using an optimized blockwise nonlocal means denoising method⁷⁷. We will combine triplane data from multiple series to reconstruct motion-corrected volumes using slice-to-volume registration⁶³. However, among multi-view images, the low-quality images with severe motion artifacts must be excluded for the best motion-correction result. Thus, we recently developed and published an automated method for **scoring the quality of fetal MR images from 0 to 1 using a deep CNN model**, which was trained with 5,051 volumes of 357 fetuses⁷⁸. The correlation between visual and automatic quality rating was highly significant (corr. coeff $r = \sim 0.8$). We will first **sort out low-score images (score < 0.2), visually recheck them for exclusion, and then perform motion correction**. The reconstruction is designed to produce 3D volume with isotropic super-resolution (0.5-0.75mm). Our motion correction has shown high performance for fetal MRI with various GWs and amounts of motion (Fig. 2). The reconstructed

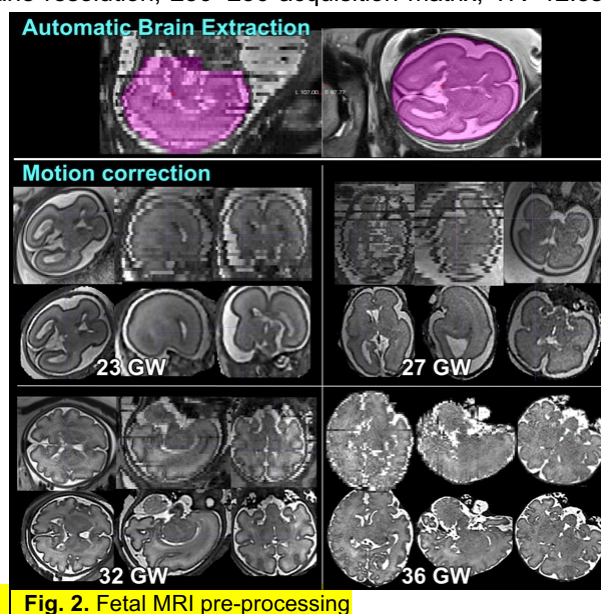


Fig. 2. Fetal MRI pre-processing

The reconstruction is designed to produce 3D volume with isotropic super-resolution (0.5-0.75mm). Our motion correction has shown high performance for fetal MRI with various GWs and amounts of motion (Fig. 2). The reconstructed

volume will be finally **inspected to see if motion artifacts and image noise still remain**. Considering the success rate of motion correction based on our extensive experience, ~85% of the data are expected to be available for deep learning training. We will align the reconstructed volume images to fetal brain volume templates (21-37 GW)⁵⁶ nearest to the GW of subjects using the FSL linear image registration tool⁷⁹.

Aim 1. Develop deep learning methods for CP, SP, and sulcal CSF segmentation from fetal brain MRI

We will develop a method to segment CP and SP and also estimate CSF in sulcal regions using 3D and 2D deep learning networks. Since the portion of sulcal CSF regions is relatively small in the whole brain, it may not be effective to segment and label all classes in one deep learning model. Thus, we will employ a **two-step approach**. First, an automatic method for CP and SP segmentation will be developed using deep CNNs. Second, **small CSF regions in deep and narrow sulci** will be estimated and detected using a patch-based CNN.

1A. Background and Preliminary Data

1A.1 Need for new segmentation methods for fetal brain MRI

In fetal brain MRI, it is more challenging to segment the thin CP and identify the boundary between CP and CSF (outer CP boundary) in highly folded sulcal regions compared to adult MRI (Fig. 3). The expectation-maximization (EM) algorithm and atlas-based segmentation method have been employed for fetal brain tissue segmentation⁸⁰⁻⁸³. They reported results of automatic CP segmentation but *utilized a small number of fetal MRIs and/or exhibited large errors* (4-16 subjects, Dice coefficient⁸⁴ [spatial overlap index between manual and automatic segmentations]=0.63-0.84 and mean surface distance [MSD] error=0.70-0.86 mm)⁸⁰⁻⁸³ (Table 1). The EM algorithm requires precise estimation of a mixture of tissue probabilities using linear and nonlinear registration between target images and a brain atlas. Likewise, atlas-based segmentation requires precise registration. **Since fetal CP and sulcal CSF regions are thin and narrow and their boundaries are blurred, accurate registration and estimation of the tissue probabilities are not easy.** Recently, deep CNN models were used for fetal CP segmentation and showed higher accuracy than prior approaches^{85,86}. However, those deep learning networks were *trained with fetal MRIs from a single-site (low generalizability) and were still not accurate enough* (Dice coefficient<0.87, MSD>0.28mm). **There has been no method for sulcal CSF segmentation in the fetal brain.** The SP is formed and grows between 13 and 30 GW and begins to slowly disappear from about 33 GW^{8,87}. The SP can be well identified on T2-weighted MR images until approximately 27 GW, but the contrast between the SP and the underlying intermediate zone is reduced between ~27 and ~31 GW^{5,6,56,88,89}. Thus, although intensity-based registration to fetal brain MRI atlas⁵⁶ was recently used for SP segmentation, large effort is still needed for manual refinement based on neuroanatomical knowledge (histological studies and images)^{5-7,55}. Since a deep learning method can utilize the spatial and other complex features as well as intensity features of brain structures, it could enhance the performance of SP segmentation despite the limited tissue contrast (See 1A.3).

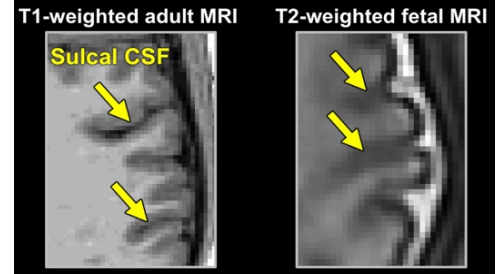


Fig. 3. Adult and fetal MRIs. Different signal intensity contrast in sulcal CSF regions

1A.2 Preliminary result: Fetal CP segmentation with multiple predictions using fully connected CNN

Recently, we published a 2D fully connected CNN for automatic CP segmentation from fetal MRI^{60,90}. We trained our deep learning model with 52 fetal brain MRIs (22.9-31.4 GW) acquired at two sites (BCH and TMC) and suggested a novel methodological approach. We adopted a *multi-view aggregation with test-time augmentation* method and a *novel hybrid loss function* to utilize 3D information and improve the segmentation accuracy (Fig. 4). Our method obtained a Dice coefficient of 0.907 and MSD error of 0.184mm and showed higher accuracy compared to prior CP segmentation methods (Table 1). **Segmentation accuracies were not influenced by different field strengths/scanners and gestational age⁶⁰.** Our hybrid loss function and multi-view aggregation significantly improved the accuracy compared to a simple Dice loss function and single-view prediction⁶⁰.

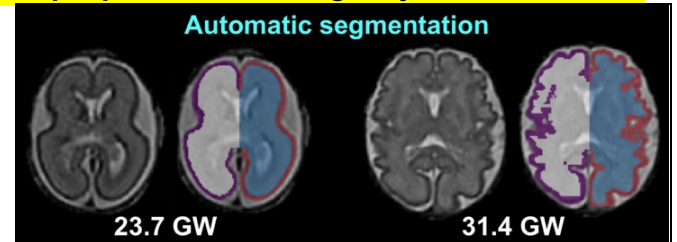


Fig. 4. Our automatic CP segmentation using a deep CNN⁶⁰

| Table 1 | Deep learning | | | EM | | | Atlas-based method |
|-----------------|---------------|-----------------------------------|-------------------------------|---------------------------------------|---------------------------------|----------------------------------|---------------------------------|
| | Our method | Khalili et al. 2019 ⁸⁶ | Dou et al. 2021 ⁸⁵ | Bach Cuadra et al. 2009 ⁸⁰ | Habas et al. 2010 ⁸¹ | Wright et al. 2014 ⁸² | Serag et al. 2012 ⁸³ |
| CP segmentation | | | | | | | |
| N | 52 | 12 | 57 | 4 | 14 | 16 | 15 |
| Dice | 0.907 ± 0.029 | 0.835 | 0.87 ± 0.06 | 0.625 ± 0.038 | 0.82 ± 0.02 | - | 0.84 ± 0.06 |
| MSD (mm) | 0.184 ± 0.063 | 0.307 | 0.28 ± 0.14 | 0.697 ± 0.079 | - | 0.864 ± 0.141 | - |

1A.3 Preliminary result: SP segmentation from T2-weighted in vivo fetal MRI

We have expertise in defining the SP region with fetal neuroanatomical knowledge on T2-weighted *in vivo* fetal MRI⁵⁻⁷ (Fig. 5a). Our experience will enable us to construct high-quality ground truth data. Using our SP

segmentation images of 20 TD fetuses (18-30 GW), we trained our same deep learning architecture described above for automatic SP segmentation. Our preliminary data show relatively good results (**Dice coefficient: ~0.85**) (Fig. 5b). Although the contrast between SP and the intermediate zone was limited after ~27 GW (29.8 GW fetus in Fig. 5) and the size of the training dataset was small, our deep learning model could detect the boundary between these two layers. These results indicate that highly accurate SP segmentation can be accomplished if we develop more robust and generalized deep learning methods.

1B. Fetal CP and SP segmentation

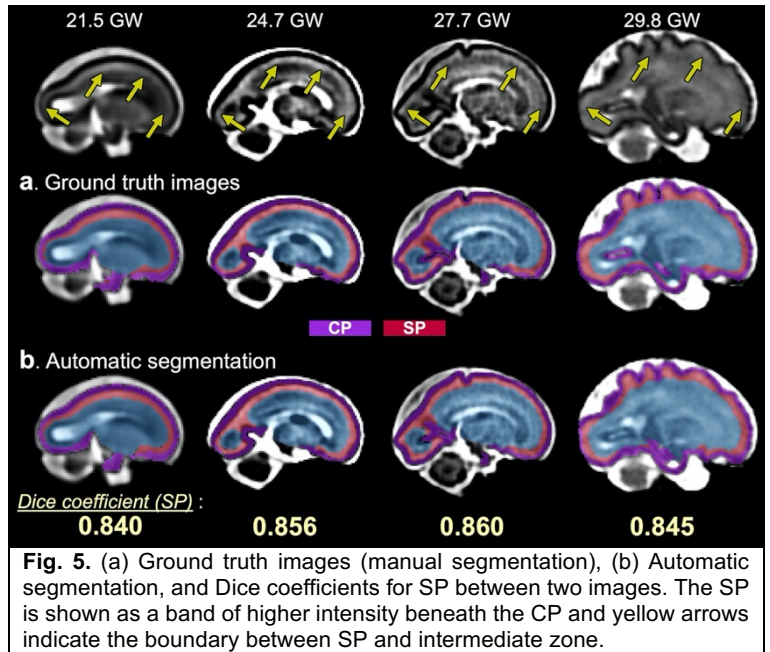
Based on our deep learning architecture used for fetal CP segmentation, we will update our deep learning tool with a larger multi-site dataset that has global applicability across a wide gestational age range (previous method: CP only, 52 fetal MRIs, 22.9-31.4 GW, proposed new method: CP, SP, and CSF, 280 fetal MRIs from three imaging sites, 18-37 GW). For the ground truth data set, CP will be automatically segmented using our tool, and segmentation errors will be manually corrected. We will also manually segment SP into left and right hemispheres separately. Two trained research assistants will perform the segmentation and error correction under the supervision of a fetal neuroanatomist (Dr. Vasung), who has expertise in CP and SP segmentation and analysis^{5,6,10}. In 15% of all cases (~40), we will test inter-rater reliability using the Dice coefficient. There will be a variability in image intensity among different sites which could affect the reliability of our model. Accordingly, we will apply two intensity normalization methods (Z-score, piecewise linear histogram matching^{91,92}) that have been widely used in many MRI studies⁹³⁻⁹⁵. **Both 3D- and 2D-based deep learning networks** will be trained with segmented images of TD fetuses. The 3D CNN has the advantages of extracting and utilizing complex 3D-based features, but it requires more parameters from 3D kernels and a larger amount of training dataset compared to the 2D CNN^{96,97}. We will employ a *biologically inspired data augmentation* strategy described below. For the 2D CNN, our *multi-view (axial, sagittal, and coronal) aggregation* approach will be used for merging features from three planes. As the SP begins to slowly disappear and can hardly be distinguished on T2-weighted MRI after 31 GW^{5,6,55,56,88}, the SP segmentation model will be trained and developed for fetuses from 18 to 31 GW.

1B.1 Network architecture

We will configure the contracting (left side) paths, expansive (right side) paths, and skip connections similar to U-Net⁷⁵, but will tune in a few points. It will consist of repeated layers of the batch normalization (BN), exponential linear units (ELU), 3×3(×3: for 3D CNN) zero-padded convolution, and a 2×2(×2) max pooling with stride=2^{98,99}. Each network layer will be divided into blocks based on the size of the feature map. Each block represents a structure in which the BN, ELU, and convolution layers are present in triplicate. The order of the layers in the block will be composed of BN, ELU, and convolution by referring to the evaluation result of the previous study¹⁰⁰. In the last layer, 1×1(×1) convolution will be used to compress the desired number of labels from the feature maps to 7 (CP, SP, and the inner volume of SP in the left and right hemispheres and background) and softmax activation will be applied to make a probability value for each label. **Additional experiments:** We will also employ other recent deep CNN architectures such as ResNet^{101,102}, DenseNet^{103,104}, Recurrent Net^{105,106}, nnU-Net¹⁰⁷, and U-Net++¹⁰⁸ and compare segmentation performance.

1B.2 Biologically inspired data augmentation

Data augmentation (simulating a large number of images from existing samples) is widely used to increase the sample size and improve the accuracy and stability of deep learning. Augmentation is often based on mathematic models – either to geometrically transform existing images (e.g., flipping, rotation, scaling)^{109,110} or, to change the image intensity^{96,97,111}, or, to learn from image textures and synthesize “fake” images¹¹². A potential limitation is that the simulated images may not emulate real variations, so may not be biologically or anatomically sound¹¹³. More recent studies used **anatomically meaningful image transformation for data augmentation** and improved segmentation accuracies compared to traditional random transformation-based augmentation¹¹⁴⁻¹¹⁶. We propose a biologically inspired augmentation framework based on atlas construction and deformable image registration, motivated by those prior studies¹¹⁴⁻¹¹⁶. Given 280 fetal subjects, we aim to simulate the same



subject's brain MRI in each of the GW from 18 to 37 GW (19 times). It will be equivalent to artificially scanning the fetal brain once every week from 18 to 37 GW, even though in reality the subject has only been scanned once. This novel augmentation will increase our sample size to 19 times the original sample size. It will then be followed by traditional augmentation (flipping along medial and horizontal sections). To implement this, we will simulate and add the “average” growth (or reverse growth) on each real image. First, we will quantify an “average brain” for each GW. For this, we will construct a virtual atlas of that week based on our extensively-validated unbiased group-wise image registration^{117,118} of all images in this age group. This is similar to our work that constructed age-specific atlases for infant brain¹¹⁹. Then, we will use our spatiotemporal deformable image registration to compute 342 ($=19 \times 18$) deformations between two virtual average brain atlases at any two GW between 18-37 GW. Each deformation T_{ij} therefore quantifies how an average brain grows (or reversely grows) from the i -th to the j -th GW (i and j between 18 and 37, $i \neq j$). This way, given a real fetal brain X_i , which was scanned on the i -th GW, we will apply the deformation T_{ij} to simulate the same fetal brain's MRI scan in the j -th GW (X_j), as if this subject was scanned on the j -th week $X_j = T_{ij}(X_i)$. **Validation:** We will compare the proposed biologically inspired augmentation with the standard geometry transformation augmentation. We expect that our augmentation method will also outperform the previous geometry transformation augmentation method.

1B.3 Hybrid loss

The Dice coefficient has been used as a loss function in the optimization of deep learning models^{86,104,120}. Recently, new loss functions that improve upon the Dice loss have been introduced^{121,122}. The Dice loss is unfavorable for relatively small structures because misclassifying a few pixels can lead to a large decrease in the coefficient¹²¹. Therefore, we will adopt a *logarithmic Dice loss (focal loss)* which focuses more on less accurate labels¹²¹. This focal loss balances between easy and hard structures to segment. Furthermore, we will use our novel *boundary Dice loss* to enhance the boundary segmentation accuracy⁶⁰. Dice loss is effective in increasing the overall overlap between ground truth and predictions but lacks segmentation accuracy for boundary areas. Thus, to increase the weight of the boundary area, we will use a *hybrid loss* by calculating the Dice loss for the boundary area and adding to the loss for the entire area. The Hausdorff distance loss was proposed to include the surface distance in the loss function¹²³. However, the computing process is complicated, and the weight compensation is difficult because the value range is different between the Dice loss and the boundary loss. We propose a morphological erosion-based boundary Dice loss⁶⁰. The proposed boundary loss is much simpler because the calculation process is similar to the whole area Dice loss, and the weight adjustment is straightforward because the range is the same as the whole area Dice loss. Our hybrid loss has shown significantly higher segmentation accuracy than a simple Dice loss function in our previous study⁶⁰.

1B.4 Multi-view aggregation for 2D CNN and a test-time augmentation for 3D and 2D CNNs

To effectively use each of the 2D planes of information from the brain MRI and enhance the accuracy of segmentation, we will use a *multi-view aggregation* approach in 2D CNN. A multi-view aggregation reflects 3D information in deep learning segmentation using 2D medical images¹⁰⁴. We will train a separate CNN for each of the three planes: axial, sagittal, and coronal planes. The predictions of each plane network will be combined into the final segmentation map (Fig. 6).

A *test-time augmentation* approach will be employed to improve the performance of segmentation¹²⁴. The test-time augmentation has been used to create various test results and combine the results to improve accuracy. Our test-time augmentation will generate multiple outputs with artificially augmented inputs such as original, horizon flip, vertical flip, and horizontal/vertical flip. We will apply the test-time augmentation on each volume and plane for 3D and 2D networks respectively, and aggregate multiple results to produce the final result (Fig. 6). In 3D CNN, the final label value on i -th voxel will be determined as $\text{argmax}_{\text{label}}(S_{\text{label},\text{vol}}(i))$. $S_{\text{label},\text{vol}}$ is the label frequency from the test-time augmentation of 3D volume. For 2D CNN, the i -th voxel will be labeled as $\text{argmax}_{\text{label}}(S_{\text{label},\text{axi}}(i) + S_{\text{label},\text{sag}}(i) + S_{\text{label},\text{cor}}(i))$, where $S_{\text{label},\text{axi}}$, $S_{\text{label},\text{sag}}$, and $S_{\text{label},\text{cor}}$ are the frequency of predicted labels using the test-time augmentation from axial, sagittal, and coronal planes, respectively.

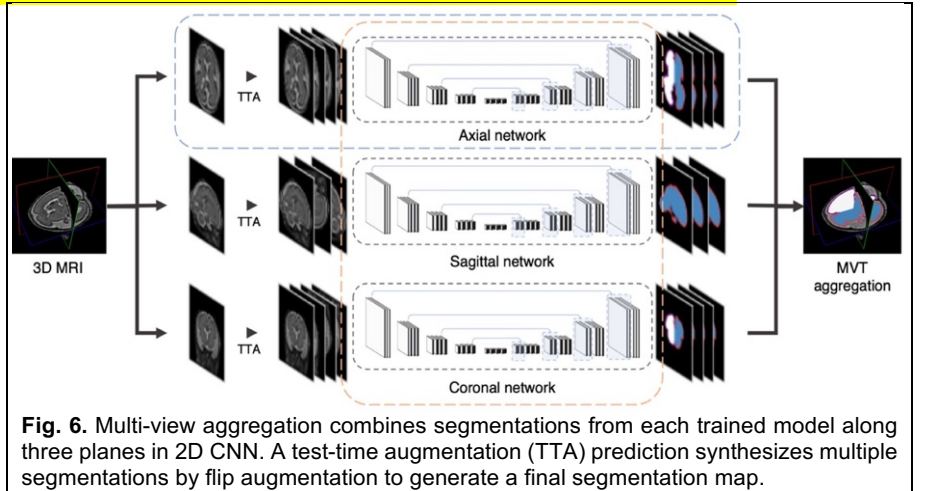


Fig. 6. Multi-view aggregation combines segmentations from each trained model along three planes in 2D CNN. A test-time augmentation (TTA) prediction synthesizes multiple segmentations by flip augmentation to generate a final segmentation map.

1C. Segmentation of deep sulcal CSF regions

Since CSF regions in deep sulci are smaller than the voxel size of fetal MRI and show partial volume effects, it is a challenging problem to generate separation between opposite sulcal banks. Previously, statistical partial volume models based on the intensity distribution of pure tissues were used to estimate the map of sulcal CSF in adult brain MRI^{25,125}. However, narrow sulcal CSF regions are too vague to be detected using partial volume estimation in fetal brain MRI (Fig. 3). In this study, after initial CP segmentation, we will design both 3D and 2D patch-based CNNs to capture the small portion of CSF in deep and narrow sulcal regions (Fig. 7). The goal of this network will be to calculate the probability (0 to 1) of CP and CSF. For a training set, the CSF outside the CP will be extracted using the dilation of the segmented CP and then sulcal CSF regions will be manually delineated.

The patch will be created only in the CP and the CSF adjacent to the outer cortical boundary. The patch-based CNN estimates the label of the central voxel of the patch by dividing the image into several patches^{58,126,127}. The 2D patch-based approach will estimate feature values from three 2D patches (axial, sagittal, and coronal) having the same center voxel using CNN and collect the estimated feature values to estimate the final label. To effectively synthesize local information and global information, two patches of different sizes will be created, and local features and global features calculated in the two patches will be combined for final information. It will allow small areas to be effectively segmented. We will apply the test-time augmentation method to the patch to generate multiple predictions and improve the segmentation performance.

1D. Training and evaluation of the CNN models

The model for CP, SP, and sulcal CSF segmentations will be evaluated using 10-fold cross-validation. In each fold of training, 10% of the training samples will be selected as a validation set. We will use a stratified sampling method to match the GW distribution between training and test datasets. To achieve segmentation accuracy without bias toward a certain age group, oversampling will be applied to training and validation datasets for a balanced uniform distribution of GW¹²⁸. The hybrid loss described above will be used for training, and deep learning will be optimized using an Adam optimizer¹²⁹. For setting the optimal network weights in each fold, we will monitor the Dice coefficient in the validation set in every epoch until there is no longer improvement of the Dice coefficient during the last 100 epochs using an early stopping function. Then the network weights at the highest Dice coefficient will be stored as the optimal network. We aim to have *Dice coefficients greater than 90% and MSD errors less than 0.18mm* between the ground truth and automatic segmentations.

External cross validation: We will perform experiments to demonstrate the generalizability of our model. We will train the CNN models using the data only from two sites/one field strength and examine how well the model performs on the data from another site/field strength using the Dice coefficient and MSD.

1E. Statistical tests and expected results

We will statistically evaluate whether the segmentation accuracies are associated with gestational age and imaging site/scanner. We will *evaluate GW-related changes of the Dice coefficient and MSD using the Pearson correlation analysis*. Segmentation accuracies will be statistically compared between different imaging sites using analysis of variance (ANOVA). Since our prior deep learning-based CP segmentation using multi-site MRI showed no significant effects of age and site/scanner (Siemens 3T vs. Philips 1.5T) on segmentation performance⁶⁰, we expect that *the accuracies will not be significantly correlated with GW and imaging sites*. We will also statistically test the effects of different methodological approaches using repeated measures ANOVA and post-hoc pairwise comparisons. The Dice coefficient and MSD error of the automatic segmentation will be compared between (1) no intensity normalization and Z-score and piecewise linear histogram matching normalizations; (2) 3D and 2D CNNs; (3) different network architectures; (4) biologically inspired augmentation and standard geometry transformation augmentation. We expect that *the use of intensity normalization and biologically inspired augmentation will significantly increase the accuracies in any deep learning model*.

1F. Potential limitations and alternative strategies

Subtle contrast between SP and intermediate zone and between CP and sulcal CSF in MRI could make accurate segmentation difficult. To increase image quality (high SNR and CNR), we will utilize an optimized blockwise nonlocal means denoising filter⁷⁷ and test several tools for slice-to-volume registration such as IRTK⁶³, BTK⁶⁴,

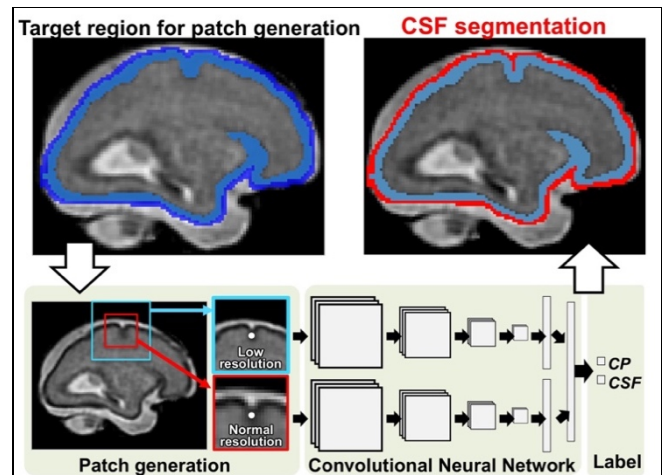


Fig. 7. Schematic representation of CSF segmentation (including sulcal regions) using a patch-based CNN.

SVRTK⁶⁵, and NiftyMIC⁶⁶ to generate the most optimal volume reconstruction. One recent study reported that echo-planar FLAIR sequence improves SP visualization in fetal MRI⁸⁸, however these are not routinely used at BCH or our collaborating sites. We will perform **FLAIR imaging for other prospective fetal projects** (Dr. Im, R01NS114087) and **test if the use of FLAIR images can improve the performance of the SP segmentation model**. The age distribution of the training dataset will be equalized by oversampling. However, if segmentation accuracy is associated with gestational age or decreased at any specific GW stage, we will give more weight to that age range in the deep learning model by increasing the training data size with more augmentation and oversampling.

Aim 2. Develop methods for outer CP and inner SP surface reconstruction and measurement of regional CP and SP thickness in the fetal brain

Inner CP surface will first be reconstructed, and then deformed outward and inward to extract both outer CP and inner SP surfaces respectively while preserving correct surface topology and vertex-wise correspondence. We will measure CP and SP thickness and perform reliability and sensitivity tests using different imaging subsets within the same subject and artificial data created by moving the boundary between CP and SP.

2A. Background and Preliminary Data

2A.1 Definition of regional correspondence between CP and SP surfaces and thickness measurement

Surface-based techniques were developed to reconstruct the inner and outer surfaces of the cerebral cortex so that the cortical thickness could be measured between the inner and outer cortical surfaces in 3D^{22-25,130}. Cortical thickness is a distance metric but there are multiple ways of defining corresponding points on the two surfaces between which that distance is to be measured. T_{link} method measures the distance between linked nodes on the inner and outer surface. The correspondence between such nodes needs to be created by the expansion of the outer surface from the inner surface, each polyhedron having the same topology and number of vertices. In adult MRI, this T_{link} approach is most sensitive in detecting altered cortical thickness compared to other methods and inherently very robust, minimizing large errors and outliers²⁸. In case of T_{near} (distance to nearest node) and T_{normal} (distance along surface normal), thickness definition could be inappropriate in highly convoluted cortical regions (Fig. 8). The proposed study will adopt

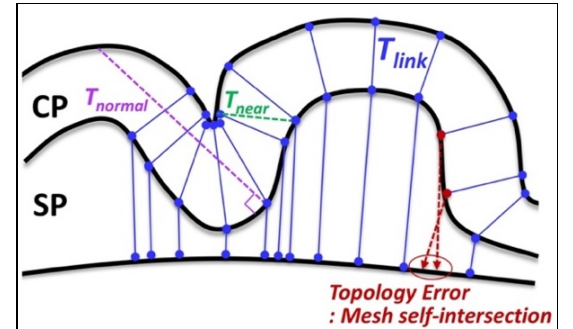


Fig. 8. Vertex correspondence between CP and SP surfaces (blue dots). T_{link} measures the distance between linked vertices. The green and purple dash lines represent improper estimations of thickness when using T_{near} and T_{normal} . The red dash line shows mesh self-intersection when deforming the inner CP surface to the inner SP boundary.

T_{link} method for measuring CP and SP thickness. Since outer CP, inner CP, and inner SP surfaces have different morphology and shape complexity, it is challenging to provide robust and reliable T_{link} -based thickness value. It is fundamental to deform one surface to another layer surface by preserving the initial surface topology and regularity of triangular meshes and having the correct deformation path and final location of vertices (Fig. 8).

2A.2 Preliminary result: Outer CP and inner SP surface reconstruction and thickness estimation

Fig. 9 shows that the initial inner CP surface was deformed to the outer CP and SP boundaries for the fetal brain at 25 GW. We computed vertex-wise CP and SP thickness using the T_{link} (Fig. 9). However, the CP and SP have a simple band-shaped structure without complex folding, and CP and SP surfaces show similar shapes in the second-trimester fetal brain. The shape and morphology of outer and inner CP and inner SP surfaces become more divergent during the third trimester. Recently, we extracted CP and SP surfaces and measured mean regional CP and SP thickness on *in utero* MRIs of 25 healthy fetuses from 20 to 33 GW⁷. The parameters for surface reconstruction were empirically modified for fetal brains of different ages, but not optimized and standardized. In the proposed study, *more robust surface deformation algorithms will be developed by systematic evaluations and accuracy tests with a large dataset, which can be generally and reliably used across a wide gestational age range*.

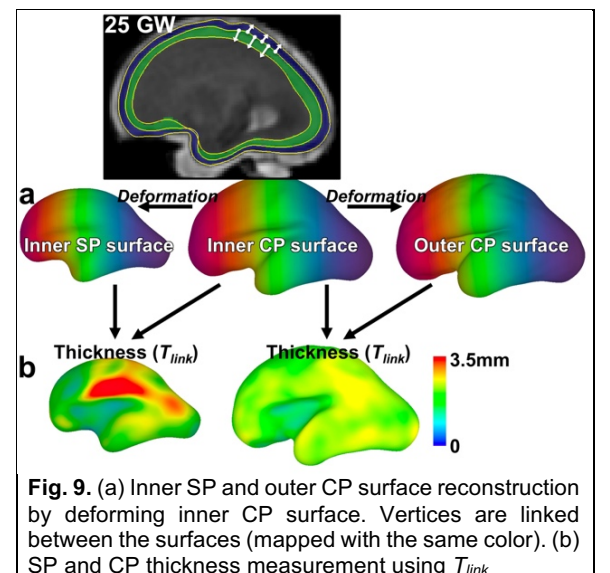


Fig. 9. (a) Inner SP and outer CP surface reconstruction by deforming inner CP surface. Vertices are linked between the surfaces (mapped with the same color). (b) SP and CP thickness measurement using T_{link}

2B. Inner and outer CP surface reconstruction

The hemispheric surface of the inner CP will be automatically tessellated as a triangular mesh and extracted (Fig. 10). The inner CP surface model will have a sphere topology based on a morphological operation. The surfaces will be geometrically smoothed to remove noisy features. The inner CP surface will be deformed under

imposed internal and external forces to reconstruct the outer CP surface. This strategy has been extensively used for outer cortical surface reconstruction^{25,51,131} because the inner cortical boundary is relatively easier to identify than the outer cortical boundary. The segmented CSF will be skeletonized to ensure correct sulcal morphology (Fig. 10). We will generate a Laplacian map containing gradients between the inner surface and CSF skeleton. This can provide a reliable expanding path from the inner cortical to the outer cortical boundary with surface topology preservation. The path of surface expansion will also guarantee the vertex-wise correspondence between inner and outer surfaces (Fig. 10).

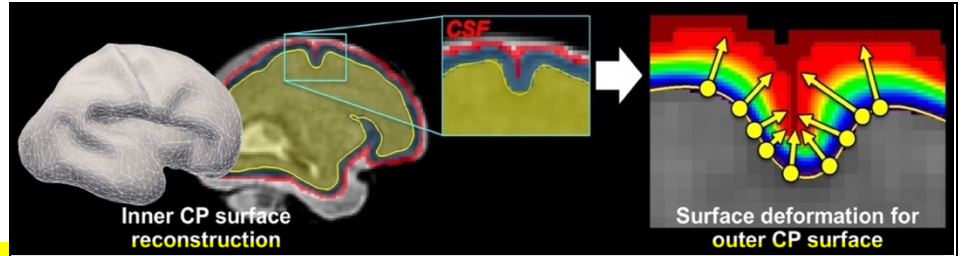


Fig. 10. Inner and outer CP surface reconstructions. The inner surface will deform based on a Laplacian map between the inner surface (yellow line) and CSF skeleton (red line)

This can provide a reliable expanding path from the inner cortical to the outer cortical boundary with surface topology preservation. The path of surface expansion will also guarantee the vertex-wise correspondence between inner and outer surfaces (Fig. 10).

In surface deformation, it is necessary to use a series of internal and external force terms for improving the accuracy of the fetal outer CP surface. In this study, we will adopt and modify the terms introduced in previous studies^{23,25}. An objective function will be defined as a weighted sum of the terms: $O(S) = \sum_k w_k T_k(S)$, where each term T_k measures some aspect of the deformable surface (S). The surface will iteratively deform itself to minimize the objective function with three terms. The first term will determine the moving distance of vertices on the gradient map at each iteration for fitting to the outer CP boundary (T_{gradient}). The second term will control the average length between vertices on the surface to regulate the mesh quality ($T_{\text{vertex-length}}$). The third term will prevent the intersection of the triangles of the surfaces ($T_{\text{intersection}}$).

For the optimization of weight parameters, we will evaluate the quality of the outer surfaces while changing the weights. First, boundary distance will be measured by an average of the minimum distances between vertices and the CSF skeleton. Smaller distance indicates surfaces are tightly fitted to the boundary of the CSF skeleton. Second, the local noise of the outer surface will be estimated by the deviation of mean curvature (H). For local noise, we will calculate the absolute sum of the deviation of mean curvature values between a vertex (v) and its

neighbors (v_j): $\frac{\sum_j |(H_v - H_{v_j}) / d(v, v_j)|}{n}$ where d is the distance between the vertex and its neighbor and n is the number of neighbors. Small noise at a vertex indicates that the local shape of the surface is smooth. Third, extracted inner and outer CP surfaces will be masked to volume images. We will isolate the voxels of the CP located between the two surfaces. We will then evaluate the Dice coefficient between the originally segmented CP and the surface-masked CP. We will explore and find the *optimal weights to extract outer CP surfaces that show the small boundary distance, low local noise, and high overlap ratio by performing several sets of experiments*. The performance of surface deformation may be influenced by the number of vertices (mesh resolution). When using the mesh with a high number of triangles, it can capture the details of boundary surface shape but suffer from the high risk of self-intersection and local noise during the surface deformation. It is also more challenging to find the balance of the weights between three force terms. Thus, we will perform the experiments with the changes in the number of triangles (327,680, 81,920, and 20,480 triangles) and find the optimal mesh resolution using the three evaluation methods. Based on vertex-wise correspondence between the inner and outer CP surfaces, we will estimate CP thickness values by measuring the distance between the linked vertices (T_{link})²⁸ (Fig. 10).

2C. Inner SP surface reconstruction

The inner CP surface will be deformed inward to the inner SP boundary based on a Laplacian map described above. Due to differences in morphology and shape complexity between inner CP and inner SP surfaces, the key point is to reconstruct the inner SP surface to have correct topology and appropriate vertex correspondence with the inner CP surface. For example, the inner CP surface is convoluted due to gyrification, but the inner SP surface shows a relatively flat shape in fetuses around 30 GW old. When we shrink the inner CP surface, the inner SP surface will have more irregular distribution and increased density of vertices (Fig. 8). Thus, in addition to T_{gradient} and $T_{\text{intersection}}$, we will use a term to consider the average position of its neighbors in determining the new location of each vertex per step (T_{stretch}), resulting in more regular edge length and smooth mesh. However, the intersection of triangles and local noise/sharp angle of the surface may still occur where vertex density is very high. To reduce that risk, the number of triangles on the mesh will be downsized (lower mesh resolution) during intermediate steps of the surface deformation. When converging to the target inner SP boundary, the triangles on the surface will be subdivided back to restore the original mesh resolution. This will also effectively reduce computing time for surface deformation. The weights for deformation force terms will be experimentally optimized using the boundary distance, local noise evaluation, and overlap ratio described above. Finally, vertex-wise SP thickness between the inner CP and inner SP surfaces will be computed using the T_{link} method²⁸.

2C.1 Evaluation with a parametric surface model

We will evaluate the reliability and accuracy of the SP surface deformation using a parametric model. A 3D sphere surface and convoluted surface modeled with 3D sine/cosine functions will be constructed and used for imitating smooth inner SP surface and folded inner CP surface (Fig. 11). Since these are mathematical parametric models, ideal vertex correspondence between the two surfaces is already determined and the distance between linked vertices can be used as the ground truth value of thickness. We will create volume images by masking the parametric surface models, run our pipeline to reconstruct and deform the surface, and generate a thickness map. Our thickness values will be compared to the ground truth values.

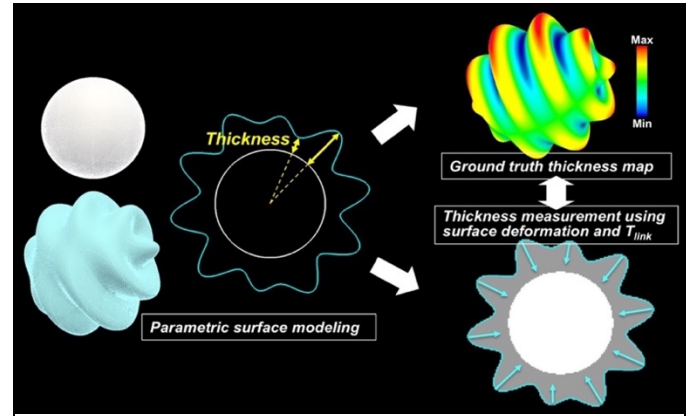


Fig. 11. Parametric models (sphere and convoluted surfaces) and ground truth thickness map. Outer convoluted surface will be deformed to the sphere volume boundary. Thickness will be measured using T_{link} and compared to the ground truth thickness.

2D. Vertex-wise definition of cortical spatial correspondence across subjects

For a statistical group analysis at a local vertex level, inter-subject vertex-wise correspondence of the surfaces will be determined using a 2D surface-based registration^{132,133}. Fetal brain surface templates are extracted from the volume templates⁵⁶. Individual cortical surfaces will be registered to the surface templates nearest to the GW of the individuals.

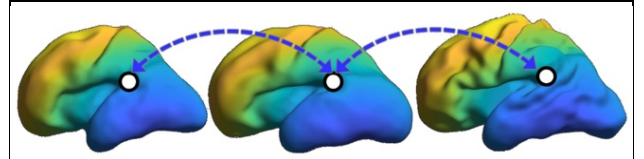


Fig. 12. Inter-subject cortical spatial correspondence

We will also align all surface templates to the template of 30 GW. As a result, we will determine the spatial correspondence between CP and SP surfaces within the same subject and across different subjects (Fig. 12). We have demonstrated high accuracy of cortical surface registration between fetal brains in our prior study⁶⁹. After cortical spatial alignment across subjects, individual CP and SP thickness maps will be transformed onto 30 GW template for further statistical analysis¹³⁴.

2E. Reliability test of fetal CP and SP thickness measurement

CP and SP segmentation and thickness measurements are based on a motion-corrected volume image that is reconstructed from raw triplane images. Multiple imaging and slice-to-volume registration for volume reconstruction may affect the measurement of fetal CP and SP thickness. Therefore, the reliability test will be performed by extracting two different CP and SP surfaces within the same subject. We will reconstruct different volumes from two randomly selected subsets (6-9 volumes for each set) of all triplane images in 50 TD fetuses (Fig. 13). We will measure CP and SP thickness from each reconstruction.

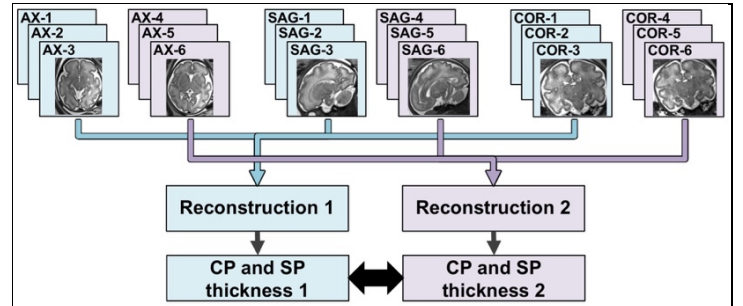


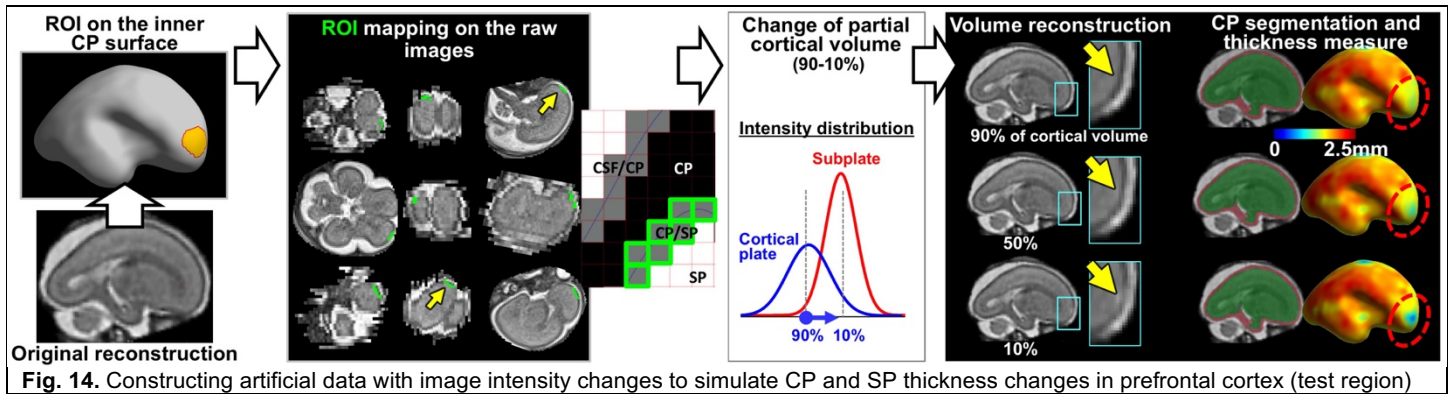
Fig. 13. CP and SP thickness comparison between two subsets of volume reconstruction within the same subject

2E.1 Statistical test and hypothesis

Inter- and intra-subject spatial correspondence of surface vertices will be determined using a 2D surface-based registration as described in 2D. A paired t -test will be performed to compare CP and SP thickness between two subsets in 50 TD fetuses. We will additionally use random-effects ANOVA to quantify the between-subject and between-subset variance of these measures. We hypothesize that no significant intra-subject differences in CP and SP thickness will be found between the subsets.

2F. Sensitivity test of fetal CP and SP thickness measurement

We will also perform a sensitivity test for CP and SP thickness using artificial data with image intensity changes designed to simulate the thickness change. Since voxel intensity is decided by a portion of brain tissues, its increase or decrease in the cortical boundary may represent cortical expansion or atrophy. Thus, we will artificially change the intensity of voxels in the CP/SP boundary (inner CP boundary) in raw triplane images. We will first define test regions (e.g. prefrontal and superior temporal regions) on the inner CP surface. These regions will be inversely mapped onto the raw images and manually edited to accurately define the inner CP boundary. Based on a partial volume estimation approach^{25,125}, we will change the portion of cortical volume from 90% to 10% in 5% intervals at each voxel belonging to the test regions (Fig. 14). We will then reconstruct a motion-corrected volume, segment CP and SP, and measure CP and SP thickness from each artificial data.



2F.1 Statistical test and hypothesis

We will estimate the changes in CP and SP thickness with the ratio of partial cortical volume using a linear mixed-effects model in 50 TD fetuses. A dependent variable will be a thickness value for each subject at each artificial data. The thickness change will be modeled with a fixed effect of the partial volume ratio and random effect accounting for within-subject correlation. We hypothesize that mean CP and SP thickness in test regions will be closely correlated with artificially adjusted cortical volume ratio. We will also determine when we can detect a significant change in CP and SP thickness from baseline (90% partial volume of CP) using a paired *t*-test.

2G. Potential limitations and alternative strategies

Fetal brains show the temporal dynamics of gyrification and CP surface shape between 18 and 37 GW^{5,10}. If our surface reconstruction is not successful for the whole age range that we propose, we will try to develop separate (two or multiple) pipelines that are optimal for specific age ranges. The weight parameters for cortical surface deformation will be optimized for different datasets with different age range.

Aim 3. Define spatiotemporal patterns of CP and SP thickness in fetal brains from 18 to 37 GW

We will model the temporal changes of regional CP and SP thickness in TD fetuses. The growth of CP and SP thickness in fetuses with cerebral abnormalities will be statistically compared to the growth of TD fetuses.

3A. Background and Preliminary Data

3A.1 Hypotheses for alterations in CP and SP thickness in PMG and ACC

Disruption in complex neural development and cortical malformations may result in altered CP thickness in the fetal stage. **PMG is a malformation of cortical development** in which the process of normal cortical development is interrupted during the stages of neuronal migration and cortical organization; the result is a disorganized cortex with multiple small gyri. PMG shows abnormally thin cortical layer structures and decreased neuronal populations and thickness of the cortex in histopathological studies^{21,42,43}. A previous MRI study also found cortical thickness reduction in adolescent patients with PMG³⁹. The SP is composed of “waiting” thalamocortical and cortico-cortical fibers, so its thickness is probably to be associated with the amount of cortico-cortical input and connections^{8,16-19}. **ACC shows a failure to develop the large bundle of fibers that connect the cerebral hemisphere**. Since there is a serious loss of interhemispheric cortical-cortical connections in ACC, it may result in reduced SP thickness¹⁷. In addition, due to the sequential ingrowth of fibers into the SP, followed by the waiting period within the SP, and the final relocation to the CP, the volumes of the SP and CP in the same areas are positively correlated¹⁰. Experimental work in animals also showed that the ablation of SP led to the disruption of the CP architecture¹³⁵. Therefore, reduced SP thickness may also cause a decrease in the CP thickness in ACC.

3A.2 Preliminary result: Disorganized cortical folding patterns in PMG and ACC

We have previously examined global sulcal folding patterns (the global patterns of positional arrangement, number, and size of sulcal folds) and revealed **disorganized sulcal patterns in fetuses and children with PMG and ACC**¹⁻⁴. In the human brain, cortical thickness strongly impacts the brain folding structure¹³⁶⁻¹³⁸. A recent study suggests that spatially variant and inhomogeneous cortical thickness plays an important role in forming human-specific stereotyped cortical folding patterns¹³⁹. Hence, *disrupted patterns of sulcal folds may be related to the altered spatial pattern of cortical thickness in PMG and ACC*.

3B. Harmonization of cortical thickness measurements across sites/scanners

Multi-site MRI data will be used to fully characterize the temporal pattern of CP and SP thickness. Although we **normalize image intensity and develop the segmentation method generalized for multi-site data**, it may still lead to non-biological variations in MRI metrics. We will **control multi-site/scanner effects using a recent technique, ComBat**^{140,141}. ComBat is a reliable tool to statistically harmonize multi-site imaging data. It has been widely used for the multi-site/scanner analyses of cross-sectional and longitudinal cortical thickness, leading to

increased statistical power compared to using non-harmonized data or using data with scanner included as a covariate^{140,142,143}. Combat was also used for multi-site diffusion and functional imaging data harmonization^{141,144}.

3C. Spatiotemporal patterns of CP and SP thickness in TD fetuses

Basic statistics, such as range (minimum and maximum), mean, and standard deviation of vertex-wise CP and SP thickness will be defined at each GW. The temporal changes of CP and SP thickness will be modeled with mathematical functions ($y = f(x)$, $x = \text{GW}$, $y = \text{CP or SP thickness}$). For CP thickness, both linear and nonlinear functions will be fitted to estimate the trend in the changes of the thickness as a function of GW. Quadratic ($b_1(x - a_1)^2 + a_2$), exponential ($a_1 e^{b_1 x} + a_2$), logistic ($a_1 + \frac{a_2}{1 + e^{-b_1(x - b_2)}}$), and Gompertz functions ($a_1 e^{-e^{-b_1(x - b_2)}} + a_2$) will be used for the nonlinear modeling, which have been previously adopted to study fetal brain growth rate (b_1)^{82,145-147}. The goodness of fit will be evaluated using R^2 value to find the most suitable function. If the logistic or Gompertz function (sigmoid function) proves most suitable for modeling CP thickness, we will also examine the GW where the maximum growth rate occurs (b_2). SP increases with age between 13 and 30 GW. Then, the SP zone diminishes and appears to be a thin band beneath the CP in the late fetal stage and disappears during the early postnatal period. A quadratic function was shown to be the best fit for the relationship between the SP volume and GW in postmortem human fetal brains¹⁰. Thus, we will employ the quadratic function ($-b_1(x - a_1)^2 + a_2$) for the fitting of SP thickness change and analyze the growth rate of SP thickness (b_1) and the maximum SP thickness (a_2). We expect that *the parameters for the growth rate of CP and SP thickness (b_1), the GW of maximum CP growth rate (b_2), and the maximum SP thickness (a_2) will be variable across different cortical areas.*

3D. Statistical analysis of alterations in CP and SP thickness in fetuses with cerebral abnormalities

CP and SP thickness at each vertex will be compared between each cerebral malformation (PMG and ACC) and TD groups. Maternal age, sex, and gestational age will be statistically matched between the groups. The same linear or nonlinear function will be used for modeling CP and SP thickness growth in the group of cerebral abnormalities. We will compare the parameters of CP/SP growth rate (b_1) between groups using an asymptotic test ($z = \Delta b_1 \div (\Sigma \text{SE}^2(b_1))^{1/2}$). If the logistic or Gompertz function is used, we will additionally compare the parameter of the maximum growth GW (b_2). The maximum thickness of the SP (a_2) will also be compared between the groups. We hypothesize that *the parameters for the growth rate of CP and SP thickness, the GW of maximum CP growth rate, and/or the maximum SP thickness will be significantly altered in fetuses with cerebral abnormalities compared to TD fetuses.*

3D.1 Power analysis

We will employ a 5% family-wise Type I error rate for comparing growth parameters between TD ($n=280$) and each of the abnormal groups separately, following the guideline of Glantz and Slinker¹⁴⁸. We have 80% power to detect a difference of 0.37σ from the PMG group ($n=70$), where σ is the underlying standard deviation of the growth parameter and 0.43σ from the ACC group ($n=50$).

3E. Interpretation of results and consideration of sex as a biological variable and multiple comparisons

Our results will be compared with previous neuroradiological and histological studies and biologically interpreted by a fetal neuroradiologist and neuroanatomist. We will consider the consistency and discrepancy between previous studies and our findings. Region-specific alterations of CP and SP thickness in PMG and ACC will be interpreted based on neural developmental events and associated with other brain structural anomalies.

In TD fetuses, the growth pattern of CP and SP thickness will be compared *between male and female subjects* using the same statistical model analysis. Because of the possibility of sex-specific variations in malformations of brain development, we will also *observe if the sex effect is significant in the thickness growth models in fetuses with PMG and ACC.*

For all statistical tests, post-hoc corrections for multiple vertex-based comparisons will be performed using the Holm step-down procedure¹⁴⁹ and false discovery rate (FDR) control at a q -value of $0.05^{150,151}$.

3F. Potential limitations and alternative strategies

Since our segmentation tool will be trained based on TD fetal brain MRIs, there may be an error of CP and SP segmentation and thickness measurement in fetal brains with cerebral abnormalities. Any segmentation errors will be manually corrected by trained experts who have knowledge of fetal neuroanatomy, and intra- and inter-rater reliability for the corrections will be tested using the Dice coefficient. For a future study, we will train and develop a segmentation method including fetuses with brain abnormalities as well as TD fetuses to improve the segmentation accuracy for abnormal fetal brains, which could facilitate more clinical applications.

4. Distribution and sharing of computational pipeline and thickness group maps

We will distribute a fully automated pipeline and group maps of TD CP and SP thickness at each GW using our **web-based platform, ChRIS, and Github repository**^{67,152-154}. See more details in the "Resource Sharing Plan".

BIBLIOGRAPHY & REFERENCES

1. Im K, Guimaraes A, Kim Y, et al. Quantitative Folding Pattern Analysis of Early Primary Sulci in Human Fetuses with Brain Abnormalities. *AJNR Am J Neuroradiol*. 2017;38(7):1449-1455. PubMed PMID: 28522661; PMCID: PMC5509490.
2. Im K, Pienaar R, Paldino MJ, Gaab N, Galaburda AM, Grant PE. Quantification and discrimination of abnormal sulcal patterns in polymicrogyria. *Cereb Cortex*. 2013;23(12):3007-3015. PubMed PMID: 22989584; PMCID: PMC3888213.
3. Tarui T, Madan N, Farhat N, et al. Disorganized Patterns of Sulcal Position in Fetal Brains with Agenesis of Corpus Callosum. *Cereb Cortex*. 2018;28(9):3192-3203. PubMed PMID: 30124828; PMCID: PMC6095213.
4. Vasung L, Yun HJ, Feldman HA, Grant PE, Im K. An Atypical Sulcal Pattern in Children with Disorders of the Corpus Callosum and Its Relation to Behavioral Outcomes. *Cereb Cortex*. 2020;30(9):4790-4799. PubMed PMID: 32307538; PMCID: PMC7391268.
5. Vasung L, Rollins CK, Yun HJ, et al. Quantitative In vivo MRI Assessment of Structural Asymmetries and Sexual Dimorphism of Transient Fetal Compartments in the Human Brain. *Cereb Cortex*. 2020;30(3):1752-1767. PubMed PMID: 31602456; PMCID: PMC7132947.
6. Vasung L, Rollins CK, Velasco-Annis C, et al. Spatiotemporal Differences in the Regional Cortical Plate and Subplate Volume Growth during Fetal Development. *Cereb Cortex*. 2020;30(8):4438-4453. PubMed PMID: 32147720; PMCID: PMC7325717.
7. Vasung L, Zhao C, Barkovich M, et al. Association between quantitative MR markers of cortical reorganization and gene expression during human prenatal brain development. *Cereb Cortex*. 2021;Accepted.
8. Kostovic I, Rakic P. Developmental history of the transient subplate zone in the visual and somatosensory cortex of the macaque monkey and human brain. *J Comp Neurol*. 1990;297(3):441-470. PubMed PMID: 2398142.
9. Bystron I, Blakemore C, Rakic P. Development of the human cerebral cortex: Boulder Committee revisited. *Nat Rev Neurosci*. 2008;9(2):110-122. PubMed PMID: 18209730.
10. Vasung L, Lepage C, Rados M, et al. Quantitative and Qualitative Analysis of Transient Fetal Compartments during Prenatal Human Brain Development. *Front Neuroanat*. 2016;10:11. PubMed PMID: 26941612; PMCID: PMC4764715.
11. Jha SC, Xia K, Schmitt JE, et al. Genetic influences on neonatal cortical thickness and surface area. *Hum Brain Mapp*. 2018;39(12):4998-5013. PubMed PMID: 30144223; PMCID: PMC6218288.
12. Jha SC, Xia K, Ahn M, et al. Environmental Influences on Infant Cortical Thickness and Surface Area. *Cereb Cortex*. 2019;29(3):1139-1149. PubMed PMID: 29420697; PMCID: PMC6373689.
13. Rakic P. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends Neurosci*. 1995;18(9):383-388. PubMed PMID: 7482803.
14. Rakic P. Evolution of the neocortex: a perspective from developmental biology. *Nat Rev Neurosci*. 2009;10(10):724-735. PubMed PMID: 19763105; PMCID: PMC2913577.
15. Golden JA, Hyman BT. Development of the superior temporal neocortex is anomalous in trisomy 21. *J Neuropathol Exp Neurol*. 1994;53(5):513-520. PubMed PMID: 8083693.
16. Judas M, Sedmak G, Kostovic I. The significance of the subplate for evolution and developmental plasticity of the human brain. *Front Hum Neurosci*. 2013;7:423. PubMed PMID: 23935575; PMCID: PMC3731572.
17. Corbett-Detig J, Habas PA, Scott JA, et al. 3D global and regional patterns of human fetal subplate growth determined in utero. *Brain Struct Funct*. 2011;215(3-4):255-263. PubMed PMID: 21046152; PMCID: PMC3041913.
18. Kostovic I, Judas M, Rados M, Hrabac P. Laminar organization of the human fetal cerebrum revealed by histochemical markers and magnetic resonance imaging. *Cereb Cortex*. 2002;12(5):536-544. PubMed PMID: 11950771.
19. McQuillen PS, Ferriero DM. Perinatal subplate neuron injury: implications for cortical development and plasticity. *Brain Pathol*. 2005;15(3):250-260. PubMed PMID: 16196392.
20. Barkovich AJ, Guerrini R, Kuzniecky RI, Jackson GD, Dobyns WB. A developmental and genetic classification for malformations of cortical development: update 2012. *Brain*. 2012;135(Pt 5):1348-1369. PubMed PMID: 22427329; PMCID: PMC3338922.

21. Judkins AR, Martinez D, Ferreira P, Dobyns WB, Golden JA. Polymicrogyria includes fusion of the molecular layer and decreased neuronal populations but normal cortical laminar organization. *J Neuropathol Exp Neurol*. 2011;70(6):438-443. PubMed PMID: 21572338; PMCID: PMC3113653.
22. Kabani N, Le Goualher G, MacDonald D, Evans AC. Measurement of cortical thickness using an automated 3-D algorithm: a validation study. *Neuroimage*. 2001;13(2):375-380. PubMed PMID: 11162277.
23. MacDonald D, Kabani N, Avis D, Evans AC. Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *Neuroimage*. 2000;12(3):340-356. PubMed PMID: 10944416.
24. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A*. 2000;97(20):11050-11055. PubMed PMID: 10984517; PMCID: PMC27146.
25. Kim JS, Singh V, Lee JK, et al. Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *Neuroimage*. 2005;27(1):210-221. PubMed PMID: 15896981.
26. Han X, Jovicich J, Salat D, et al. Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. *Neuroimage*. 2006;32(1):180-194. PubMed PMID: 16651008.
27. Iscan Z, Jin TB, Kendrick A, et al. Test-retest reliability of freesurfer measurements within and between sites: Effects of visual approval process. *Hum Brain Mapp*. 2015;36(9):3472-3485. PubMed PMID: 26033168; PMCID: PMC4545736.
28. Lerch JP, Evans AC. Cortical thickness analysis examined through power analysis and a population simulation. *Neuroimage*. 2005;24(1):163-173. PubMed PMID: 15588607.
29. Wonderlick JS, Ziegler DA, Hosseini-Varnamkhasti P, et al. Reliability of MRI-derived cortical and subcortical morphometric measures: effects of pulse sequence, voxel geometry, and parallel imaging. *Neuroimage*. 2009;44(4):1324-1333. PubMed PMID: 19038349; PMCID: PMC2739882.
30. Shaw P, Kabani NJ, Lerch JP, et al. Neurodevelopmental trajectories of the human cerebral cortex. *J Neurosci*. 2008;28(14):3586-3594. PubMed PMID: 18385317; PMCID: PMC6671079.
31. Im K, Lee JM, Lee J, et al. Gender difference analysis of cortical thickness in healthy young adults with surface-based methods. *Neuroimage*. 2006;31(1):31-38. PubMed PMID: 16426865.
32. Shaw P, Greenstein D, Lerch J, et al. Intellectual ability and cortical development in children and adolescents. *Nature*. 2006;440(7084):676-679. PubMed PMID: 16572172.
33. Hazlett HC, Gu H, Munsell BC, et al. Early brain development in infants at high risk for autism spectrum disorder. *Nature*. 2017;542(7641):348-351. PubMed PMID: 28202961; PMCID: PMC5336143.
34. Hayes JP, Logue MW, Sadeh N, et al. Mild traumatic brain injury is associated with reduced cortical thickness in those at risk for Alzheimer's disease. *Brain*. 2017;140(3):813-825. PubMed PMID: 28077398; PMCID: PMC6075586.
35. Sailer M, Fischl B, Salat D, et al. Focal thinning of the cerebral cortex in multiple sclerosis. *Brain*. 2003;126(Pt 8):1734-1744. PubMed PMID: 12805100.
36. Bullmore E. Cortical Thickness and Connectivity in Schizophrenia. *Am J Psychiatry*. 2019;176(7):505-506. PubMed PMID: 31256617.
37. Im K, Lee JM, Seo SW, et al. Variations in cortical thickness with dementia severity in Alzheimer's disease. *Neurosci Lett*. 2008;436(2):227-231. PubMed PMID: 18400396.
38. Lin JJ, Salamon N, Lee AD, et al. Reduced neocortical thickness and complexity mapped in mesial temporal lobe epilepsy with hippocampal sclerosis. *Cereb Cortex*. 2007;17(9):2007-2018. PubMed PMID: 17088374.
39. Oliveira PP, Jr., Valente KD, Shergill SS, Leite Cda C, Amaro E, Jr. Cortical thickness reduction of normal appearing cortex in patients with polymicrogyria. *J Neuroimaging*. 2010;20(1):46-52. PubMed PMID: 19453835.
40. Li G, Wang L, Shi F, et al. Cortical thickness and surface area in neonates at high risk for schizophrenia. *Brain Struct Funct*. 2016;221(1):447-461. PubMed PMID: 25362539; PMCID: PMC4452433.
41. Wang F, Lian C, Wu Z, et al. Developmental topography of cortical thickness during infancy. *Proc Natl Acad Sci U S A*. 2019;116(32):15855-15860. PubMed PMID: 31332010; PMCID: PMC6689940.
42. Desikan RS, Barkovich AJ. Malformations of cortical development. *Ann Neurol*. 2016;80(6):797-810. PubMed PMID: 27862206; PMCID: PMC5177533.

43. Golden JA, Harding BN. Cortical malformations: unfolding polymicrogyria. *Nat Rev Neurol*. 2010;6(9):471-472. PubMed PMID: 20811463.
44. Kostovic I, Judas M. Transient patterns of cortical lamination during prenatal life: do they have implications for treatment? *Neurosci Biobehav Rev*. 2007;31(8):1157-1168. PubMed PMID: 17586047.
45. Carlo CN, Stevens CF. Structural uniformity of neocortex, revisited. *Proc Natl Acad Sci U S A*. 2013;110(4):1488-1493. PubMed PMID: 23297199; PMCID: PMC3557031.
46. la Fougere C, Grant S, Kostikov A, et al. Where in-vivo imaging meets cytoarchitectonics: the relationship between cortical thickness and neuronal density measured with high-resolution [18F]flumazenil-PET. *Neuroimage*. 2011;56(3):951-960. PubMed PMID: 21073964.
47. Pletikos M, Sousa AM, Sedmak G, et al. Temporal specification and bilaterality of human neocortical topographic gene expression. *Neuron*. 2014;81(2):321-332. PubMed PMID: 24373884; PMCID: PMC3931000.
48. Miller JA, Ding SL, Sunkin SM, et al. Transcriptional landscape of the prenatal human brain. *Nature*. 2014;508(7495):199-206. PubMed PMID: 24695229; PMCID: PMC4105188.
49. Kang HJ, Kawasawa YI, Cheng F, et al. Spatio-temporal transcriptome of the human brain. *Nature*. 2011;478(7370):483-489. PubMed PMID: 22031440; PMCID: PMC3566780.
50. Morton PD, Korotcova L, Lewis BK, et al. Abnormal neurogenesis and cortical growth in congenital heart disease. *Sci Transl Med*. 2017;9(374). PubMed PMID: 28123074; PMCID: PMC5467873.
51. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*. 1999;9(2):179-194. PubMed PMID: 9931268.
52. Fischl B. FreeSurfer. *Neuroimage*. 2012;62(2):774-781. PubMed PMID: 22248573; PMCID: PMC3685476.
53. Gholipour A, Estroff JA, Barnewolt CE, et al. Fetal MRI: A Technical Update with Educational Aspirations. *Concepts Magn Reson Part A Bridg Educ Res*. 2014;43(6):237-266. PubMed PMID: 26225129; PMCID: PMC4515352.
54. Wang X, Pettersson DR, Studholme C, Kroenke CD. Characterization of Laminar Zones in the Mid-Gestation Primate Brain with Magnetic Resonance Imaging and Histological Methods. *Front Neuroanat*. 2015;9:147. PubMed PMID: 26635541; PMCID: PMC4656822.
55. Rollins CK, Ortinau CM, Stopp C, et al. Regional Brain Growth Trajectories in Fetuses with Congenital Heart Disease. *Ann Neurol*. 2021;89(1):143-157. PubMed PMID: 33084086.
56. Gholipour A, Rollins CK, Velasco-Annis C, et al. A normative spatiotemporal MRI atlas of the fetal brain for automatic segmentation and analysis of early brain growth. *Sci Rep*. 2017;7(1):476. PubMed PMID: 28352082; PMCID: PMC5428658.
57. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444. PubMed PMID: 26017442.
58. Moeskops P, Viergever MA, Mendrik AM, de Vries LS, Benders MJ, Isgum I. Automatic Segmentation of MR Brain Images With a Convolutional Neural Network. *IEEE Trans Med Imaging*. 2016;35(5):1252-1261. PubMed PMID: 27046893.
59. Wachinger C, Reuter M, Klein T. DeepNAT: Deep convolutional neural network for segmenting neuroanatomy. *Neuroimage*. 2018;170:434-445. PubMed PMID: 28223187; PMCID: PMC5563492.
60. Hong J, Yun HJ, Park G, et al. Fetal Cortical Plate Segmentation Using Fully Convolutional Networks With Multiple Plane Aggregation. *Front Neurosci*. 2020;14:591683. PubMed PMID: 33343286; PMCID: PMC7738480.
61. Gholipour A, Estroff JA, Warfield SK. Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain MRI. *IEEE Trans Med Imaging*. 2010;29(10):1739-1758. PubMed PMID: 20529730; PMCID: PMC3694441.
62. Kim K, Habas PA, Rousseau F, Glenn OA, Barkovich AJ, Studholme C. Intersection based motion correction of multislice MRI for 3-D in utero fetal brain image formation. *IEEE Trans Med Imaging*. 2010;29(1):146-158. PubMed PMID: 19744911; PMCID: PMC3328314.
63. Kuklisova-Murgasova M, Quaghebeur G, Rutherford MA, Hajnal JV, Schnabel JA. Reconstruction of fetal brain MRI with intensity matching and complete outlier removal. *Med Image Anal*. 2012;16(8):1550-1564. PubMed PMID: 22939612.
64. Rousseau F, Oubel E, Pontabry J, et al. BTK: an open-source toolkit for fetal brain MR image processing. *Comput Methods Programs Biomed*. 2013;109(1):65-73. PubMed PMID: 23036854; PMCID: PMC3508300.

65. Uus A, Zhang T, Jackson LH, et al. Deformable Slice-to-Volume Registration for Motion Correction of Fetal Body and Placenta MRI. *IEEE Trans Med Imaging*. 2020. PubMed PMID: 32086200.
66. Ebner M, Wang G, Li W, et al. An automated framework for localization, segmentation and super-resolution reconstruction of fetal brain MRI. *Neuroimage*. 2020;206:116324. PubMed PMID: 31704293; PMCID: PMC7103783.
67. <https://chrisproject.org>.
68. Im K, Jo HJ, Mangin JF, Evans AC, Kim SI, Lee JM. Spatial distribution of deep sulcal landmarks and hemispherical asymmetry on the cortical surface. *Cereb Cortex*. 2010;20(3):602-611. PubMed PMID: 19561060.
69. Yun HJ, Chung AW, Vasung L, et al. Automatic labeling of cortical sulci for the human fetal brain based on spatio-temporal information of gyrification. *Neuroimage*. 2019;188:473-482. PubMed PMID: 30553042; PMCID: PMC6452886.
70. Yun HJ, Im K, Jin-Ju Y, Yoon U, Lee JM. Automated sulcal depth measurement on cortical surface reflecting geometrical properties of sulci. *PLoS One*. 2013;8(2):e55977. PubMed PMID: 23418488; PMCID: PMC3572156.
71. Ortinau CM, Rollins CK, Gholipour A, et al. Early-Emerging Sulcal Patterns Are Atypical in Fetuses with Congenital Heart Disease. *Cereb Cortex*. 2019;29(8):3605-3616. PubMed PMID: 30272144; PMCID: PMC6644862.
72. Ad-Dab'bagh Y, Einarson D, Lyttelton O, et al. The CIVET Image-Processing Environment: A Fully Automated Comprehensive Pipeline for Anatomical Neuroimaging Research. In *Proceedings of the 12th Annual Meeting of the Organization for Human Brain Mapping, Florence, Italy*. 2006.
73. Yamashita Y, Namimoto T, Abe Y, et al. MR imaging of the fetus by a HASTE sequence. *AJR American journal of roentgenology*. 1997;168(2):513-519. PubMed PMID: 9016238.
74. <https://github.com/FNNDSC/pl-fetal-brain-mask#usage>.
75. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. 2015:234-241.
76. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging*. 2010;29(6):1310-1320. PubMed PMID: 20378467; PMCID: PMC3071855.
77. Coupe P, Yger P, Prima S, Hellier P, Kervrann C, Barillot C. An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Trans Med Imaging*. 2008;27(4):425-441. PubMed PMID: 18390341; PMCID: PMC2881565.
78. <https://github.com/ilegorreta/Automatic-Fetal-Brain-Quality-Assessment-Tool>.
79. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*. 2002;17(2):825-841. PubMed PMID: 12377157.
80. Bach Cuadra M, Schaer M, Andre A, Guibaud L, Eliez S, Thiran J-P. Brain tissue segmentation of fetal MR images. *International Conference on Medical Image Computing and Computer Assisted Intervention: Workshop on Image Analysis for Developing Brain*. 2009.
81. Habas PA, Kim K, Rousseau F, Glenn OA, Barkovich AJ, Studholme C. Atlas-based segmentation of developing tissues in the human brain with quantitative validation in young fetuses. *Hum Brain Mapp*. 2010;31(9):1348-1358. PubMed PMID: 20108226; PMCID: PMC3306251.
82. Wright R, Kyriakopoulou V, Ledig C, et al. Automatic quantification of normal cortical folding patterns from fetal brain MRI. *Neuroimage*. 2014;91:21-32. PubMed PMID: 24473102.
83. Serag A, Kyriakopoulou V, Rutherford MA, et al. A Multi-channel 4D Probabilistic Atlas of the Developing Brain: Application to Fetuses and Neonates. *Annals of the BMVA*. 2012;3:1-14.
84. Dice LR. Measures of the amount of ecologic association between species. *Ecology*. 1945;26(297-302).
85. Dou H, Karimi D, Rollins CK, et al. A Deep Attentive Convolutional Neural Network for Automatic Cortical Plate Segmentation in Fetal MRI. *IEEE Trans Med Imaging*. 2020;PP. PubMed PMID: 33351755.
86. Khalili N, Lessmann N, Turk E, et al. Automatic brain tissue segmentation in fetal MRI using convolutional neural networks. *Magn Reson Imaging*. 2019;64:77-89. PubMed PMID: 31181246.
87. Kostovic I, Judas M, Sedmak G. Developmental history of the subplate zone, subplate neurons and interstitial white matter neurons: relevance for schizophrenia. *Int J Dev Neurosci*. 2011;29(3):193-205. PubMed PMID: 20883772.

88. Diogo MC, Prayer D, Gruber GM, et al. Echo-planar FLAIR Sequence Improves Subplate Visualization in Fetal MRI of the Brain. *Radiology*. 2019;292(1):159-169. PubMed PMID: 31084478.
89. Widjaja E, Geibprasert S, Mahmoodabadi SZ, Blaser S, Brown NE, Shannon P. Alteration of human fetal subplate layer and intermediate zone during normal development on MR and diffusion tensor imaging. *AJNR Am J Neuroradiol*. 2010;31(6):1091-1099. PubMed PMID: 20075102.
90. https://github.com/jwhong1125/fetal_CP_segmentation.
91. Nyul LG, Udupa JK, Zhang X. New variants of a method of MRI scale standardization. *IEEE Trans Med Imaging*. 2000;19(2):143-150. PubMed PMID: 10784285.
92. Shah M, Xiao Y, Subbanna N, et al. Evaluating intensity normalization on MRIs of human brain with multiple sclerosis. *Med Image Anal*. 2011;15(2):267-282. PubMed PMID: 21233004.
93. Carre A, Klausner G, Edjlali M, et al. Standardization of brain MR images across machines and protocols: bridging the gap for MRI-based radiomics. *Sci Rep*. 2020;10(1):12340. PubMed PMID: 32704007; PMCID: PMC7378556.
94. Martensson G, Ferreira D, Granberg T, et al. The reliability of a deep learning model in clinical out-of-distribution MRI data: A multicohort study. *Med Image Anal*. 2020;66:101714. PubMed PMID: 33007638.
95. Reinhold JC, Dewey BE, Carass A, Prince JL. Evaluating the Impact of Intensity Normalization on MR Image Synthesis. *Proc SPIE Int Soc Opt Eng*. 2019;10949. PubMed PMID: 31551645; PMCID: PMC6758567.
96. Bashyam VM, Erus G, Doshi J, et al. MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14 468 individuals worldwide. *Brain*. 2020;143(7):2312-2324. PubMed PMID: 32591831; PMCID: PMC7364766.
97. Fu J, Yang Y, Singhrao K, et al. Deep learning approaches using 2D and 3D convolutional neural networks for generating male pelvic synthetic computed tomography from magnetic resonance imaging. *Med Phys*. 2019;46(9):3788-3798. PubMed PMID: 31220353.
98. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *32nd International Conference on Machine Learning, ICML 2015 International Machine Learning Society (IMLS)*. 2015:448-456.
99. Clevert DA, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (ELUs). *4th International Conference on Learning Representations, ICLR 2016*. 2016.
100. He K, Zhang X, Ren S, Sun J. Identity Mappings in Deep Residual Networks. *Lecture Notes in Computer Science, Computer Vision – ECCV 2016*. 2016:630-645.
101. Chen H, Dou Q, Yu L, Qin J, Heng PA. VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images. *Neuroimage*. 2018;170:446-455. PubMed PMID: 28445774.
102. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016:770-778.
103. Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely Connected Convolutional Networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*. 2017:2261-2269.
104. Guha Roy A, Conjeti S, Navab N, Wachinger C, Alzheimer's Disease Neuroimaging I. QuickNAT: A fully convolutional network for quick and accurate segmentation of neuroanatomy. *Neuroimage*. 2019;186:713-727. PubMed PMID: 30502445.
105. Alom MZ, Yakopcic C, Hasan M, Taha TM, Asari VK. Recurrent residual U-Net for medical image segmentation. *J Med Imaging (Bellingham)*. 2019;6(1):014006. PubMed PMID: 30944843; PMCID: PMC6435980.
106. Liang M, Hu X. Recurrent convolutional neural network for object recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015:3367-3375.
107. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods*. 2021;18(2):203-211. PubMed PMID: 33288961.
108. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support (2018)*. 2018;11045:3-11. PubMed PMID: 32613207; PMCID: PMC7329239.
109. Shen D, Wu G, Suk HI. Deep Learning in Medical Image Analysis. *Annu Rev Biomed Eng*. 2017;19:221-248. PubMed PMID: 28301734; PMCID: PMC5479722.

110. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal.* 2017;42:60-88. PubMed PMID: 28778026.
111. Nalepa J, Marcinkiewicz M, Kawulok M. Data Augmentation for Brain-Tumor Segmentation: A Review. *Front Comput Neurosci.* 2019;13:83. PubMed PMID: 31920608; PMCID: PMC6917660.
112. Sandfort V, Yan K, Pickhardt PJ, Summers RM. Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. *Sci Rep.* 2019;9(1):16884. PubMed PMID: 31729403; PMCID: PMC6858365.
113. Eaton-Rosen Z, Bragman F, Ourselin S, Cardoso MJ. Improving data augmentation for medical image segmentation. *International Conference on Medical Image with Deep Learning.* 2018.
114. Zhao A, Balakrishnan G, Durand F, Guttag JV, Dalca AV. Data augmentation using learned transformations for one-shot medical image segmentation. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* 2019.
115. Silva M, Garcia K, Sudre CH, Bass C, Cardoso MJ, Robinson E. Biomechanical modelling of brain atrophy through deep learning. *arXiv:201207596.* 2020.
116. Zhao Y, Rhee DJ, Cardenas C, Court LE, Yang J. Training deep-learning segmentation models from severely limited data. *Med Phys.* 2021. PubMed PMID: 33474727.
117. Ou Y, Akbari H, Bilello M, Da X, Davatzikos C. Comparative evaluation of registration algorithms in different brain databases with varying difficulty: results and insights. *IEEE Trans Med Imaging.* 2014;33(10):2039-2065. PubMed PMID: 24951685; PMCID: PMC4371548.
118. Ou Y, Sotiras A, Paragios N, Davatzikos C. DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting. *Med Image Anal.* 2011;15(4):622-639. PubMed PMID: 20688559; PMCID: PMC3012150.
119. Ou Y, Zollei L, Retzepi K, et al. Using clinically acquired MRI to construct age-specific ADC atlases: Quantifying spatiotemporal ADC changes from birth to 6-year old. *Hum Brain Mapp.* 2017;38(6):3052-3068. PubMed PMID: 28371107; PMCID: PMC5426959.
120. Wang G, Li W, Ourselin S, Vercauteren T. Automatic Brain Tumor Segmentation Using Cascaded Anisotropic Convolutional Neural Networks. *In International MICCAI brainlesion workshop.* 2017:178-190.
121. Wong KC, Moradi M, Tang H, Syeda-Mahmood T. 3D Segmentation with Exponential Logarithmic Loss for Highly Unbalanced Object Sizes. *International Conference on Medical Image Computing and Computer-Assisted Intervention.* 2018:612-619.
122. Sudre CH, Li W, Vercauteren T, Ourselin S, Jorge Cardoso M. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).* 2017:240-248.
123. Karimi D, Salcudean SE. Reducing the Hausdorff Distance in Medical Image Segmentation With Convolutional Neural Networks. *IEEE Trans Med Imaging.* 2020;39(2):499-513. PubMed PMID: 31329113.
124. Wang G, Li W, Aertsen M, Deprest J, Ourselin S, Vercauteren T. Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. *Neurocomputing.* 2019;335:34-45. PubMed PMID: 31595105; PMCID: PMC6783308.
125. Tohka J, Zijdenbos A, Evans A. Fast and robust parameter estimation for statistical partial volume models in brain MRI. *Neuroimage.* 2004;23(1):84-97. PubMed PMID: 15325355.
126. Kushibar K, Valverde S, Gonzalez-Villa S, et al. Automated sub-cortical brain structure segmentation combining spatial and deep convolutional features. *Med Image Anal.* 2018;48:177-186. PubMed PMID: 29935442.
127. Kamnitsas K, Ledig C, Newcombe VFJ, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal.* 2017;36:61-78. PubMed PMID: 27865153.
128. Feng X, Lipton ZC, Yang J, et al. Estimating brain age based on a uniform healthy population with deep learning and structural magnetic resonance imaging. *Neurobiol Aging.* 2020;91:15-25. PubMed PMID: 32305781.
129. Kingma DP, Ba JL. Adam: a method for stochastic optimization. *3rd Int Conf on Learn Represent, ICLR 2015 - Conf Track Proc, San Diego, CA.* 2015.
130. Jones SE, Buchbinder BR, Aharon I. Three-dimensional mapping of cortical thickness using Laplace's equation. *Hum Brain Mapp.* 2000;11(1):12-32. PubMed PMID: 10997850.

131. Liu T, Nie J, Tarokh A, Guo L, Wong ST. Reconstruction of central cortical surface from brain MRI images: method and application. *Neuroimage*. 2008;40(3):991-1002. PubMed PMID: 18289879; PMCID: PMC2505350.
132. Lyttelton O, Boucher M, Robbins S, Evans A. An unbiased iterative group registration template for cortical surface analysis. *Neuroimage*. 2007;34(4):1535-1544. PubMed PMID: 17188895.
133. Robbins S, Evans AC, Collins DL, Whitesides S. Tuning and comparing spatial normalization methods. *Med Image Anal*. 2004;8(3):311-323. PubMed PMID: 15450225.
134. Chung MK, Robbins SM, Dalton KM, Davidson RJ, Alexander AL, Evans AC. Cortical thickness analysis in autism with heat kernel smoothing. *Neuroimage*. 2005;25(4):1256-1265. PubMed PMID: 15850743.
135. Kanold PO, Luhmann HJ. The subplate and early cortical circuits. *Annu Rev Neurosci*. 2010;33:23-48. PubMed PMID: 20201645.
136. Heuer K, Toro R. Role of mechanical morphogenesis in the development and evolution of the neocortex. *Phys Life Rev*. 2019;31:233-239. PubMed PMID: 30738760.
137. Mota B, Herculano-Houzel S. BRAIN STRUCTURE. Cortical folding scales universally with surface area and thickness, not number of neurons. *Science*. 2015;349(6243):74-77. PubMed PMID: 26138976.
138. Richman DP, Stewart RM, Hutchinson JW, Caviness VS, Jr. Mechanical model of brain convolutional development. *Science*. 1975;189(4196):18-21. PubMed PMID: 1135626.
139. Costa Campos LD, Hornung R, Gompfer G, Elgeti J, Caspers S. The role of thickness inhomogeneities in hierarchical cortical folding. *Neuroimage*. 2021;231:117779. PubMed PMID: 33548459.
140. Fortin JP, Cullen N, Sheline YI, et al. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage*. 2018;167:104-120. PubMed PMID: 29155184; PMCID: PMC5845848.
141. Fortin JP, Parker D, Tunc B, et al. Harmonization of multi-site diffusion tensor imaging data. *Neuroimage*. 2017;161:149-170. PubMed PMID: 28826946; PMCID: PMC5736019.
142. Beer JC, Tustison NJ, Cook PA, et al. Longitudinal ComBat: A method for harmonizing longitudinal multi-scanner imaging data. *Neuroimage*. 2020;220:117129. PubMed PMID: 32640273.
143. Radua J, Vieta E, Shinohara R, et al. Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. *Neuroimage*. 2020;218:116956. PubMed PMID: 32470572; PMCID: PMC7524039.
144. Yu M, Linn KA, Cook PA, et al. Statistical harmonization corrects site effects in functional connectivity measurements from multi-site fMRI data. *Hum Brain Mapp*. 2018;39(11):4213-4227. PubMed PMID: 29962049; PMCID: PMC6179920.
145. Xia J, Wang F, Benkarim OM, et al. Fetal cortical surface atlas parcellation based on growth patterns. *Hum Brain Mapp*. 2019;40(13):3881-3899. PubMed PMID: 31106942.
146. Makropoulos A, Aljabar P, Wright R, et al. Regional growth and atlasing of the developing human brain. *Neuroimage*. 2016;125:456-478. PubMed PMID: 26499811; PMCID: PMC4692521.
147. Tarui T, Im K, Madan N, et al. Quantitative MRI Analyses of Regional Brain Growth in Living Fetuses with Down Syndrome. *Cereb Cortex*. 2019. PubMed PMID: 31264685.
148. Glantz SA, Slinker BK. Primer of Applied Regression and Analysis of Variance. *McGraw-Hill, New York*. 1990:303.
149. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Statist*. 1979;6(2):65-70.
150. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*. 1995;57:289-300.
151. Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*. 2002;15(4):870-878. PubMed PMID: 11906227.
152. <https://github.com/FNNDSC/cookiecutter-chrisapp>.
153. <https://chrisstore.co>.
154. <https://app.chrisproject.org>.