Check for updates

METHOD ARTICLE

REVISED A spatio-temporal atlas of the developing fetal brain with spina bifida aperta

[version 2; peer review: 2 approved]

Lucas Fidon[1], Elizabeth Viola[1], Nada Mufti[1,2], Anna L. David[2,3], Andrew Melbourne[1], Philippe Demaerel[4], Sébastien Ourselin[1], Tom Vercauteren[1], Jan Deprest[2-4], Michael Aertsen[4]

[1]School of Biomedical Engineering & Imaging Sciences, King's College London, London, SE1 7EU, UK
[2]Elizabeth Garrett Anderson Institute for Women's Health, University College London, London, WC1E 6DB, UK
[3]Department of Obstetrics and Gynaecology, University Hospitals Leuven, 3000 Leuven, Belgium
[4]Department of Radiology, University Hospitals Leuven, 3000 Leuven, Belgium

v2 **First published:** 15 Oct 2021, **1**:123
https://doi.org/10.12688/openreseurope.13914.1
Latest published: 31 Aug 2022, **1**:123
https://doi.org/10.12688/openreseurope.13914.2

Abstract

Background: Spina bifida aperta (SBA) is a birth defect associated with severe anatomical changes in the developing fetal brain. Brain magnetic resonance imaging (MRI) atlases are popular tools for studying neuropathology in the brain anatomy, but previous fetal brain MRI atlases have focused on the normal fetal brain. We aimed to develop a spatio-temporal fetal brain MRI atlas for SBA.
Methods: We developed a semi-automatic computational method to compute the first spatio-temporal fetal brain MRI atlas for SBA. We used 90 MRIs of fetuses with SBA with gestational ages ranging from 21 to 35 weeks. Isotropic and motion-free 3D reconstructed MRIs were obtained for all the examinations. We propose a protocol for the annotation of anatomical landmarks in brain 3D MRI of fetuses with SBA with the aim of making spatial alignment of abnormal fetal brain MRIs more robust. In addition, we propose a weighted generalized Procrustes method based on the anatomical landmarks for the initialization of the atlas. The proposed weighted generalized Procrustes can handle temporal regularization and missing annotations. After initialization, the atlas is refined iteratively using non-linear image registration based on the image intensity and the anatomical land-marks. A semi-automatic method is used to obtain a parcellation of our fetal brain atlas into eight tissue types: white matter, ventricular system, cerebellum, extra-axial cerebrospinal fluid, cortical gray matter, deep gray matter, brainstem, and corpus callosum.
Results: An intra-rater variability analysis suggests that the seven anatomical land-marks are sufficiently reliable. We find that the proposed atlas outperforms a normal fetal brain atlas for the automatic segmentation of brain 3D MRI of fetuses with SBA.

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 31 Aug 2022		✓ view
version 1 15 Oct 2021	✓ view	? view

1. **Andras Jakab**, University Children's Hospital Zürich, Zürich, Switzerland

2. **Guillaume Auzias**, Aix-Marseille University, Marseille, France

Any reports and responses or comments on the article can be found at the end of the article.

Conclusions: We make publicly available a spatio-temporal fetal brain MRI atlas for SBA, available here: https://doi.org/10.7303/syn25887675. This atlas can support future research on automatic segmentation methods for brain 3D MRI of fetuses with SBA.

Keywords
Fetal brain atlas, spina bifida aperta, fetal brain development, fetal brain T2-weighted MRI, anatomical landmarks, spatio-temporal atlas, segmentation, super resolution and reconstruction



This article is included in the Medical Sciences gateway.



This article is included in the Horizon 2020 gateway.



This article is included in the Marie-Sklodowska-Curie Actions (MSCA) gateway.



This article is included in the Brain Dysfunction and Degeneration collection.

Corresponding author: Lucas Fidon (lucas.fidon@kcl.ac.uk)

Author roles: Fidon L: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation; **Viola E**: Data Curation, Formal Analysis, Investigation, Methodology, Validation, Visualization, Writing – Original Draft Preparation; **Mufti N**: Resources, Writing – Review & Editing; **David AL**: Resources, Writing – Review & Editing; **Melbourne A**: Resources, Writing – Review & Editing; **Demaerel P**: Data Curation; **Ourselin S**: Project Administration, Resources, Supervision; **Vercauteren T**: Conceptualization, Project Administration, Resources, Supervision, Writing – Review & Editing; **Deprest J**: Conceptualization, Project Administration, Resources, Supervision; **Aertsen M**: Conceptualization, Data Curation, Resources, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No [765148]. This work was also supported by core and project funding from the Wellcome Trust and the Engineering and Physical Sciences Research Council (EPSRC) [WT203148/Z/16/Z; NS/A000049/1; WT101957; NS/A000027/1]. ALD is supported by the National Institute for Health University College London Hospitals Biomedical Research Centre.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2022 Fidon L *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Fidon L, Viola E, Mufti N *et al.* **A spatio-temporal atlas of the developing fetal brain with spina bifida aperta [version 2; peer review: 2 approved]** Open Research Europe 2022, **1**:123 https://doi.org/10.12688/openreseurope.13914.2

First published: 15 Oct 2021, **1**:123 https://doi.org/10.12688/openreseurope.13914.1

REVISED **Amendments from Version 1**

The validation cohort has been expanded to include the fetal brain 3D MRIs from the FeTA dataset releases 1 and 2. Two new figures (Figure 9 and Figure 10) have been added to show boxplots of the Dice score and Hausdorff distance for each week of gestational age and each tissue type. The proposed fetal brain atlas has been updated in the public synapse repository with improved segmentations for the atlases at weeks 33 and 34. A figure showing the distribution per gestational age of the genders of the fetuses whose 3D MRIs were used to compute the atlas have been added in Figure 1B. We have added Figure 14 to illustrate the variability in the topology of the extra-axial cerebrospinal fluid at 26 weeks of gestation in operated fetuses with spina bifida aperta. We have also clarified some parts of the method, the discussion, and the limitations with the help of the reviewers' feedback.

Any further responses from the reviewers can be found at the end of the article

Plain language summary

Approximately five per 10,000 babies born in Europe suffer from spina bifida aperta (SBA). SBA is a birth defect that occurs when the spinal column of the fetus fails to close during the first month of pregnancy. SBA can impact the development of the fetal brain, resulting in lifelong disabilities such as cognitive impairment, difficulties with mobility, and a reduced life expectancy. The effect of SBA on the development of the fetal brain is complex and is not yet fully understood. Developing our understanding of SBA is fundamental to improving diagnosis and management for babies born with this condition. Fetal brain atlases are maps of the development of the fetal brain during the pregnancy. Such atlases allow researchers to perform measurements of the fetal brain anatomy and to study the development of the fetal brain in a large population. However, current fetal brain atlases only correspond to normal fetal brain development. In this work, we developed the first atlas of the developing brain in fetuses with SBA between 21 weeks and 34 weeks of gestation. This condition-specific atlas will allow us to perform more accurate measurements in fetuses with SBA. The atlas is created from 90 Magnetic Resonance Imaging (MRI) scans taken of fetuses with SBA in the womb, a technique which allows the fetal brain to be visualised in 3D and in high resolution. The period 21–34 weeks of the development of the fetal brain in SBA is of particular interest because surgery performed while the baby is still in the womb is currently completed prior to 26 weeks of gestation. The proposed atlas could therefore support research on the effect of the surgery on the fetal brain anatomy.

1 Introduction

Spina bifida aperta (SBA) is the most prevalent fetal brain defect with approximately five per 10,000 live births in Europe[1]. It occurs when the neural tube fails to close in the first four weeks after conception. Most cases of SBA are accompanied by severe anatomical brain abnormalities[2] with enlargement of the ventricles and a type II Chiari malformation being most prevalent. The Chiari malformation type II is characterized by a small posterior fossa and hindbrain herniation in which the medulla, cerebellum, and fourth ventricule are displaced caudally into the spinal canal[3]. The corpus callosum of fetuses with SBA is also abnormal[2,4] and has been found to be significantly smaller for fetuses with SBA than for normal fetuses[4–6]. In postnatal life, children and adults with spina bifida aperta are known to have also smaller hippocampus[7], abnormal cortical thickness and gyrification[8,9], and smaller deep grey matter volume and total brain volume[10,11]. In a small pilot study, it has been observed that fetal brain volume and shape is different after spina bifida repair compared to controls[12].

For all those reasons the anatomy of the brain of fetuses with SBA differs from the normal fetal brain anatomy. In addition, the mechanisms underlying those anatomical brain abnormalities remain incompletely understood[13].

Brain atlases are used to study common trends and variations in the brain anatomy of a population. They provide a model of a population of brain magnetic resonance images (MRIs) that represents the average brain anatomy of a population, allow the comparison of measurements in a cohort study, and can be used for the automatic segmentation of brain MRIs[14–17]. Atlases can also be used to measure variability in the brain anatomy of an individual as compared to the model supposed to be representative of the whole population[14]. Age and disease specific atlases allow a more accurate model of specific populations of human brains to be obtained[18].

Previous work on fetal brain atlases has focused on age-specific atlases by proposing various spatio-temporal fetal brain MRI atlases[14,15,17,19–22]. A spatio-temporal atlas does not consist in only one average volume, but instead consists in a collection of age-specific average volumes. This allows the development of the fetal brain anatomy to be modelled. However, existing studies have only used brain MRIs of fetuses with a normal brain development, except for one study that combined fetuses with a normal brain and fetuses with lissencephaly

in the same atlas[14]. In particular, no fetal brain atlas for the developing fetal brain with SBA has been proposed in the literature.

In this work, we propose the first spatio-temporal fetal brain MRI atlas for SBA. Our atlas covers all the weeks of gestation between 21 weeks and 34 weeks. This range of gestational ages is of particular interest for SBA because it starts before the time at which in-utero surgery for SBA is currently performed[13] and covers most of the time until birth. The atlas is computed using 90 fetal brain MRIs from 37 fetuses with SBA. We hypothesise that the high variability of the brain anatomy in SBA is one of the main challenges in adapting methods developed for normal fetal brain atlases for SBA. To tackle this issue, we propose a semi-automatic method for the computation of the proposed fetal brain MRI atlas for SBA. We propose a protocol for the annotation of 11 anatomical landmarks in fetal brain 3D MRI of fetuses. Those anatomical landmarks are used in two important steps of our pipeline. The anatomical landmarks are used firstly to initialize the computation of the atlas using a weighted generalized Procrustes method and secondly to regularize the non-linear image registration of fetal brain 3D MRIs to the atlas.

We performed an intra-rater variability evaluation for the proposed landmarks using a subset of 31 3D MRIs from our cohort. Based on this evaluation, 4 anatomical landmarks were excluded and 7 were selected to help for the computation of the spatio-temporal atlas. In addition, we evaluated the automatic fetal brain segmentations computed using the proposed atlas for SBA on 40 fetal brain 3D MRIs of the publicly available FeTA dataset[23]. It contains 15 MRIs of normal fetuses and 25 MRIs of fetuses with SBA. We compared the automatic segmentations computed using our SBA atlas to the segmentations computed using a state-of-the-art normal fetal brain MRI atlas[20]. We have found that the proposed SBA atlas outperforms the normal fetal brain atlas on cases with SBA. The proposed spatio-temporal fetal brain MRI atlas for SBA is made publicly available here.

2 Materials

In this section, we describe the fetal brain MRI data used to compute the atlas and for the evaluation of automatic segmentations obtained using the atlas.

2.1 Ethics statement

The MRI data were automatically pseudonymized using the `GIFT-Cloud` data sharing platform[24] prior to using them for research.

At University Hospitals Leuven, ethical approval to use the data for research was given by the Ethics Committee University Hospitals Leuven (ethical approval S63598). A retrospective study does not fall under the Belgian law of May 7, 2004 regarding experiments on the human person. However, given the use of potentially identifying MRIs in the study, the requirements set forth in the EU Regulation 2016/679 (General Data Protection Regulation, GDPR) must be met. The sponsor of this study is University Hospitals Leuven, and University Hospitals Leuven maintains “public interest” as the legal basis for data processing. Article 14 of the GDPR mentions the information obligation of the data controller (= sponsor of the study) to the data subject whose personal data are collected. An information obligation is therefore sufficient according to GDPR, and informed consent is not legally required for the use of the MRIs for illustrative purposes. All snapshots of fetal MRIs used in our figures are based on MRIs acquired at Leuven.

At University College London Hospital (UCLH) the study was approved by the Caldicott guardian at UCLH and patient consent was not required as these images were acquired for clinical purposes and the data used retrospectively.

2.2 Spina bifida aperta cohort used to compute the spatio-temporal atlas

A total of 90 fetal brain MRI examinations from 37 fetuses were used in this work.

All the MRI examinations were performed as part of clinical routine following abnormal findings during ultrasound examination. All the fetuses in this cohort were diagnosed with spina bifida aperta at fetal ultrasound examinations. MRI scans were acquired at two surgical centers, University Hospitals Leuven and UCLH (see *Underlying data*). For each study, at least three orthogonal T2-weighted HASTE series of the fetal brain were collected on a 1.5T scanner using an echo time of 133ms, a repetition time of 1000ms, with no slice overlap nor gap, pixel size 0.39mm to 1.48mm, and slice thickness 2.50mm to 4.40mm. A radiologist attended all the acquisitions for quality control. The dataset contains longitudinal MRI examinations with up to 5 examinations

per fetus. In addition, 51 of the MRI examinations were performed after open fetal surgery performed before 26 weeks of gestation, to close the spina bifida aperta defect. The distribution of gestational ages for MRI examinations and whether they were done before or after surgery can be found in Figure 1.

2.3 Fetal brain 3D MRI used for the evaluation of automatic segmentation

For the evaluation of automatic fetal brain segmentation we have used the publicly available FeTA dataset[23,25] (first and second release).

The FeTA dataset contains 90 reconstructed 3D MRIs, including 32 MRIs of fetuses with a normal brain (gestational ages from 21 weeks to 35 weeks) and 38 MRIs of fetuses with spina bifida aperta (gestational ages from 20 weeks to 30 weeks). The others are MRIs of fetuses with other abnormalities and were therefore excluded. For all the 3D MRIs, segmentations are available for seven tissue types: white matter, ventricular system, cerebellum, extra-axial cerebrospinal fluid, cortical grey matter, deep grey matter, and brainstem.

The 40 3D MRIs and original segmentations (as provided with the FeTA dataset) were inspected by two paediatric radiologists within our institutions, MA and PD, with more than 8 years of experience in segmenting fetal brains. Corrections of the segmentations were performed[26–28] to reduce the variability against the published segmentation guidelines that was released with the FeTA dataset[23,25]. Two volumes of spina bifida aperta cases (`sub-feta007` and `sub-feta009`) were excluded because the poor quality of the 3D reconstruction did not allow to segment them reliably for the seven tissue types. This implies a total of 36 3D MRIs of spina bifida subjects were used for evaluation.

2.4 Spatio-temporal atlas for the normal developing fetal brain

For comparison to a spatio-temporal atlas of the normal developing fetal brain, we have used the publicly available spatio-temporal fetal brain atlas[20]. This atlas contains 18 3D MRIs of average normal fetal brain for gestational ages ranging from 21 weeks to 38 weeks.

3 Atlas computation method

In this section, we describe our pipeline for computing the spina bifida aperta (SBA) fetal brain 4D atlas. An overview of the pipeline can be found in Figure 2.

3.1 Data preprocessing

In this subsection, we give details about the preprocessing steps as can be found in Figure 2.

3.1.1 Automatic brain segmentation. One of the main challenges in fetal brain MRI is the motion of the fetus. To tackle this issue, MRI sequences used for fetal MRI are designed to produce multiple stacks of 2D slices rather than a 3D image. Original 2D slices typically have lower resolution, suffers from motion between

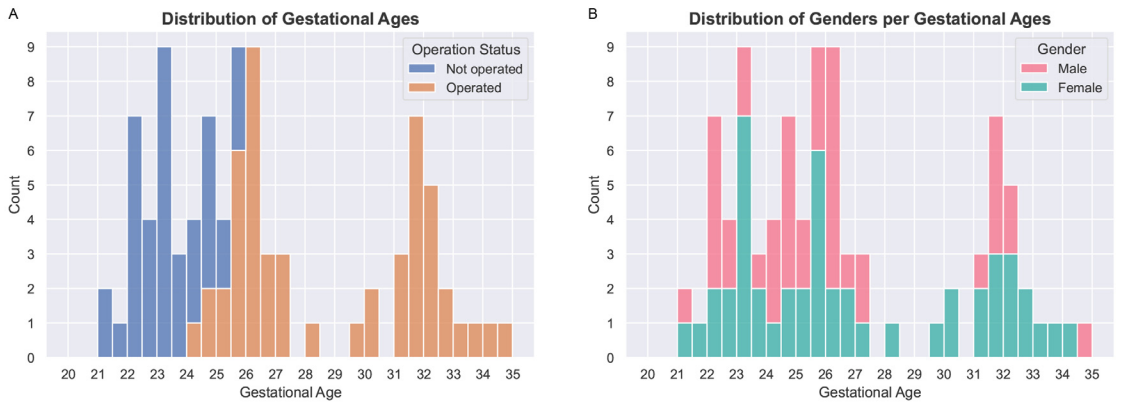


Figure 1. (**A**) Distribution of gestational ages for operated (fetal surgery) and non-operated fetal brains. The dataset used to compute the atlas contains 39 magnetic resonance imaging (MRI) examinations of non-operated fetuses and 51 MRI examinations of operated fetuses. (**B**) Distribution of genders per gestational age. We found no statistical difference between the distributions of gestational ages for males and females using a Mann-Whitney U test with a confidence level of 95%.

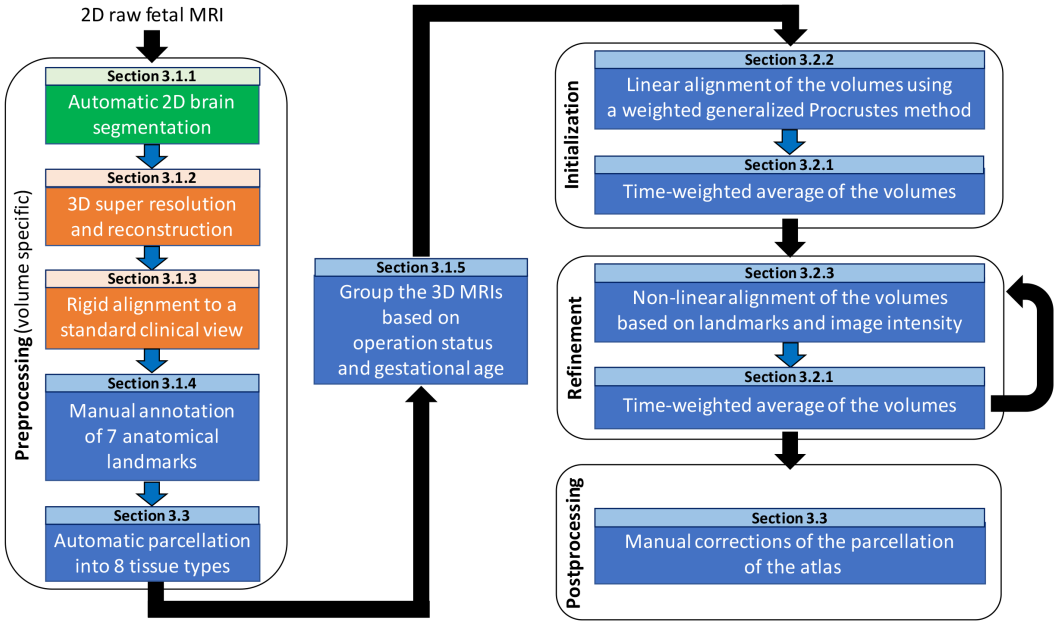


Figure 2. Overview of the spatio-temporal atlas construction pipeline. MRI: magnetic resonance imaging. In green are the steps computed using MONAIfbs[29] and in orange are the steps computed using NiftyMIC[30].

neighboring slices, motion artefact, and suboptimal cross-section[30]. Automatic segmentation of the fetal brain in the raw 2D MRI are obtained using a deep learning-based method[29]. Those brain masks are an input required by the 3D super resolution and reconstruction algorithm described below. A public implementation of the deep learning pipeline `MONAIfbs`[29], used in this study to obtain the brain masks, can be found here (main git branch, commit *bcab52a*).

3.1.2 3D super resolution and reconstruction. We use a 3D super resolution and reconstruction algorithm to improve the resolution, and remove motion between neighboring slices and motion artefacts present in the original 2D slices[30]. The output of the 3D super resolution and reconstruction algorithm[30] is a reconstructed 3D MRI of the fetal brain with an isotropic image resolution (of 0.8 mm in our case). We hypothesize that the reconstructed 3D MRI facilitates the manual delineation and annotation of the fetal brain structures as compared to the original 2D slices.

We used a state-of-the-art 3D super resolution and reconstruction algorithm[30] publicly available in the `NiftyMIC` pipeline version 0.8 with Python 3.8. The original 2D MRI slices were also corrected for bias field in the `NiftyMIC` pipeline version 0.8 using a N4 bias field correction step as implemented in SimpleITK version 1.2.4. The 3D super resolution and reconstruction algorithm[30] also combines the brain masks obtained in section Automatic brain segmentation. This results in a 3D brain mask for the 3D reconstructed MRI that is computed fully-automatically.

3.1.3 Rigid alignment to a standard clinical view. The 3D reconstructed MRI were rigidly aligned to a time-point volume of the control fetal brain 4D atlas[20] as implemented in `NiftyMIC`[30] version 0.8. All the 3D reconstructed MRIs are therefore aligned to a standard clinical view in which the axes are aligned with the axial, sagittal, and coronal planes of the fetal brain. This facilitates the manual delineation and annotation of the fetal brain structures. The target time-point in the control 4D atlas is chosen based on the brain volume computed using the automatic 3D brain mask.

3.1.4 Anatomical landmarks. Seven anatomical landmarks were manually annotated to regularize and improve the accuracy of the image registration steps used in the computation of the spina bifida 4D atlas. Details can be found in section Atlas construction.

The anatomical landmarks that were selected are: the right and left anterior horn of the lateral ventricles, the posterior tectum plate, the right and left junctions between the cerebellum and the brainstem, and the right and left deep grey matter border at the foramen of Monro. An illustration of those anatomical landmarks can be found in Figure 3.

Those landmarks include anatomical structures that have been reported to be reliably identifiable in the fetal MRI clinical research literature[31–33]. Another selection criteria was to choose landmarks that are spread over the fetal brain anatomy to efficiently support image registration. Our proposed annotation protocol can be found in Annotation potocol of anatomical landmarks for fetuses with spina bifida aperta.

The manual annotations of the 90 3D reconstructed MRIs were performed by author EV. Manual annotations of landmarks were performed using the software ITK-SNAP[34] version 3.8.0. The annotation of one volume took 12 min on average. It is worth noting, that landmarks can be missing, especially for fetal MRIs before 26 weeks of gestation.

The intra-rater reliability for the anatomical landmarks has been evaluated, as described in Section Intra-rater variability for the annotation of the anatomical landmarks. The proposed anatomical landmarks protocol also included the right and left deep grey matter border at the anterior cavum septi pellucidi line and the right and left deep grey matter border at the posterior cavum septi pellucidi line. However, those landmarks were found to be unreliable and often missing due to the high variation in shape of the cavum septi pellucidi. For this reason, those landmarks were not used for the computation of the atlas but they are present in the annotation protocol. Details can be found in Section Intra-rater variability for the annotation of the anatomical landmarks.

3.1.5 Age and operation status specific groups of 3D reconstructed MRIs. The 3D reconstructed MRIs were grouped with respect to their operation status and their gestational age. Each group of 3D reconstructed MRIs went through the atlas construction pipeline described in section Atlas construction and lead to the computation of a unique volume of our spatio-temporal atlas.

SBA surgery affects the evolution of the fetal brain anatomy[8,13,31]. Therefore, we have chosen to separate the 3D reconstructed MRIs of operated and non-operated fetuses. A group either contains only 3D reconstructed MRIs of fetuses that have been operated for SBA in-utero, or contains only 3D reconstructed MRIs of fetuses that have not been operated.

Each group is assigned with a gestational age ranging from 21 weeks to 34 weeks. Volumes are included in a group only if the gestational age at the time of the acquisition is within 9 days of the gestational age of the group. This implies that there are overlaps between groups. For example, the 24 weeks group contains the fetal brain MRIs acquired between 22 weeks + 4 days and 25 weeks + 3 days of gestation. In addition, the contribution of each volume within an age-specific group is weighted using a time-varying Gaussian kernel, as defined in the next section in (1). The value of 9 days, used above, is chosen to correspond to 3 x sigma where sigma is defined in the time-varying Gaussian kernel regression (1). The description of the cohort used can

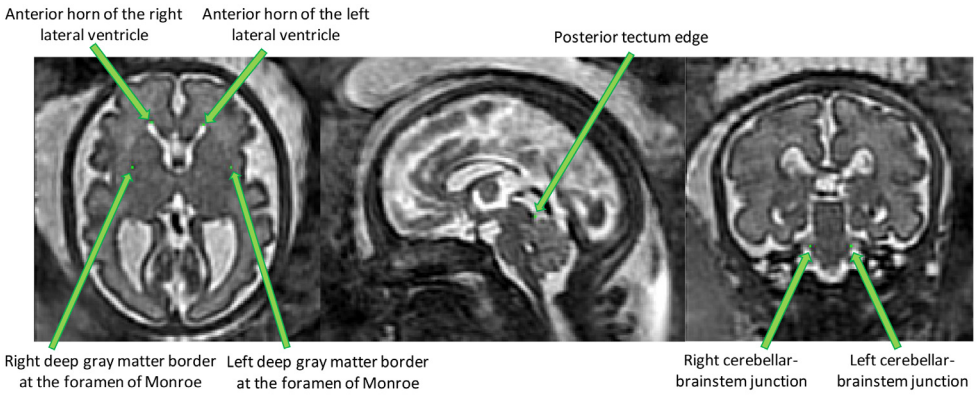


Figure 3. Overview of the proposed anatomical landmarks. Those landmarks were annotated for all the 3D reconstructed magnetic resonance imaging (MRI). They aim at improving the accuracy and the robustness of the image registration steps.

be found in section Spina bifida aperta cohort used to compute the spatio-temporal atlas and the distribution of gestation ages can be found in Figure 1. As can be seen in Figure 11 and Figure 12, groups for non-operated fetuses cover the gestational ages from 21 weeks to 25 weeks and groups for operated fetuses cover gestational ages from 25 weeks to 34 weeks.

A group is excluded if it contains less than three 3D reconstructed MRIs. In addition, we excluded a group if it did not include both 3D reconstructed MRIs with gestational ages higher and lower than the gestational age of the group. This avoids, for example, to have a group for non-operated fetuses at 26 weeks of gestation that would contain only MRIs at gestational ages 25 weeks or less.

Data augmentation: We used right-left flipping as a data augmentation to synthetically increase the amount of volumes in each group. This encourages the atlas to be symmetrical with respect to the central sagittal plane. Right-left flipping has been used in several previous studies on brain MRI atlases[35,36]. Imposing symmetry between right and left hemispheres of the atlas volumes aims at reducing potential biases in the cohort used to compute the atlas. In addition, it allows to use the atlas for the study of asymmetry between right and left hemispheres[36]. Asymmetry between brain hemispheres for normal fetuses has been described as well as the role of hemispheric asymmetry in isolated corpus callosum agenesis[37,38]. To the best of our knowledge, hemispheric asymmetry has not been studied yet in SBA.

3.2 Atlas construction

In this section we describe the different steps for the computation of the spina bifida atlas as can be seen in the Initialization and Refinement boxes of the pipeline overview in Figure 2.

3.2.1 Time-weighted average of the volumes. In this section, we describe the method to average the intensity of 3D reconstructed MRIs after spatial alignment. As described in section Age and operation status specific groups of 3D reconstructed MRIs, data are grouped with respect to their operation status and gestational age. After aligning spatially all the 3D reconstructed MRIs of a group, we average their image intensity to obtain an average fetal brain MRI for the group.

Time-weighted average: To reflect the gestational age associated with each group, we used a time-weighted average. The weight for the volume i is defined using a Gaussian kernel as follow[17]

$$w_i = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2} \left(\frac{GA_i - GA_{target}}{\sigma} \right)^2 \right) \tag{1}$$

where GA_{target} is the gestational age of the group and GA_i is the gestational age of volume i. The standard deviation value is set to $\sigma = 3$ days. We have chosen the value $\sigma = 3$ days so that an interval $[-\sigma, \sigma]$ covers approximately one week which is the time unit for the atlas.

In addition, we average each image and its symmetric by right-left flipping to impose to the average volume to be exactly symmetric with respect to the central sagittal plane. This is performed in addition to the data augmentation described in section Age and operation status specific groups of 3D reconstructed MRIs.

Formally, let $\{I_i\}_{i=1}^N$ be a set of N co-registered 3D reconstructed MRIs to average. The weighted average is computed as

$$I_{average} = \frac{1}{2N} \sum_{i=1}^{N} w_i (I_i + S(I_i)) \tag{2}$$

where S is the operator that computes the symmetric of a volume with respect to the central sagittal plane.

Preprocessing: Before averaging, we transform the intensity of each volume linearly to set the mean (resp. the standard deviation) of the image intensity inside the brain mask to 2000 (resp. 500). Those values were set to approximate the intensity profile of a spatio-temporal fetal brain atlas of normal fetuses[20].

3.2.2 Time-weighted generalized Procrustes. In this section, we describe the optimization method that we used for the joint initial linear alignment of the volumes in a group of 3D reconstructed MRIs. This method is based on a weighted generalized Procrustes method and uses only the anatomical landmarks. Especially, note that the image intensity is not used.

Generalized Procrustes methods[39] aims at matching simultaneously n configurations of landmarks using linear spatial transformations. Generalized Procrustes methods (without constraints) can be defined as optimization problems of the form[39,40]

$$\min_{\{M_i,t_i\}} \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{K}\left\| M_i x_{i,k} + t_i - \frac{1}{n}\sum_{j=1}^{n}(M_j x_{j,k} + t_j)\right\|^2 \tag{3}$$

where n is the number of samples, K is the number of landmarks, $x_{i,k}$ is the vector of coordinates for the landmark k of sample i, t_i is the translation for the sample i, and M_i is the linear transformation for the sample i. In this work we restrict the linear transformations M_i to be anisotropic scaling transformations.

However, for the computation of the spina bifida atlas we have to take into account that landmarks can be missing for some samples. We also would like to weight differently the samples based on their gestational age alike what is done for the weighted average of the 3D reconstructed MRIs in section Time-weighted average of the volumes.

In this work, we introduce weights in the generalized Procrustes methods. A weight of zeros represents a missing landmark for a sample. The proposed weighted generalized Procrustes method corresponds to the optimization problem

$$\min_{\{M_i,t_i\}} \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{K} w_{i,k}\left\| M_i x_{i,k} + t_i - \frac{\sum_{j=1}^{n} w_{j,k}(M_j x_{j,k} + t_j)}{\sum_{j=1}^{n} w_{j,k}}\right\|^2 \tag{4}$$

where $w_{i,k} \geq 0$ is the weight for the landmark k of sample i. For landmark k, sample i of gestational age GA_i, and the target gestational age GA_{target}, we propose to define the weight $w_{i,k}$ as

$$w_{i,k} = \begin{cases} 0 & \text{if landmark k is missing for sample i} \\ 0 & \text{if } |GA_i - GA_{target}| > 3\sigma \\ \dfrac{1}{\sqrt{2\pi}\sigma}\exp\left(-\dfrac{1}{2}\left(\dfrac{GA_i - GA_{target}}{\sigma}\right)^2\right) & \text{otherwise} \end{cases} \tag{5}$$

The standard deviation value is $\sigma = 3$ days.

We assume that every landmark was annotated at least once in each group. As a result, $\forall k, \sum_{j=1}^{n} w_{j,k} > 0$ and the fractions used in (4) are well defined.

In general, the optimization problem (3) admits an infinity of solutions, including the trivial solution that send all the landmarks to the origin. To tackle this issue, constraints on the size of the system are added[39,40]. The optimization problem (4) suffers from the same under-specification problem. We therefore choose to constrain the center of mass of the barycenter of the system and the size of the system because it is the most intuitive approach. This leads to the optimization problem

$$\begin{aligned} &\min_{\{M_i,t_i\},\{g_k\}} \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{K} w_{i,k}\left\| M_i x_{i,k} + t_i - g_k\right\|^2 \\ &\text{s.t.} \quad \frac{1}{K}\sum_{k=1}^{K} g_k = \frac{1}{K}\sum_{k=1}\frac{\sum_{j=1}^{n} w_{j,k} x_{j,k}}{\sum_{j=1}^{n} w_{j,k}} \\ &\text{and} \quad \frac{1}{K}\sum_{k=1}^{K}\left\| g_k - \frac{1}{K}\sum_{l=1}^{K} g_l\right\|^2 = \frac{1}{K}\sum_{k=1}\left\| \frac{\sum_{j=1}^{n} w_{j,k} x_{j,k}}{\sum_{j=1}^{n} w_{j,k}} - \frac{1}{K}\sum_{l=1}\frac{\sum_{j=1}^{n} w_{j,l} x_{j,l}}{\sum_{j=1}^{n} w_{j,l}}\right\|^2 \end{aligned} \tag{6}$$

This optimization problem can be solved efficiently using an alternating least squares approach[40].

3.2.3 Non-linear image registration. In this section, we describe the non-linear image registration method that we used for the refinement step of the 4D atlas as can be seen in Figure 2. In the refinement step, intermediate atlas MRI volumes have already been computed for all time points. The goal of this step is to improve the image sharpness of the intermediate atlas MRI volumes by registering all the 3D reconstructed MRIs to the intermediate MRI volumes and computing new weighted average volumes using the method described in section Time-weighted average of the volumes.

We used `NiftyReg`[41] to perform non-linear image registration using image intensity and the anatomical landmarks.

The non-linear image registration optimization problem is the following

$$\min_{\Theta} \mathscr{L}(I_{subject}, I_{atlas}, \phi(\Theta)) + R(\Theta) \tag{7}$$

where $I_{subject}$ is the 3D reconstructed MRI to be aligned to the 3D atlas time point I_{atlas} and $\phi(\Theta)$ is a spatial transformation parameterized by cubic B-splines of parameters Θ.

The regularization term R is a linear combination of the bending energy[42] (BE) and the linear energy[42] (LE) regularization functions applied to $\phi(\Theta)$

$$R(\Theta) = \alpha_{BE} BE(\phi(\Theta)) + \alpha_{LE} LE(\phi(\Theta)) \tag{8}$$

with α_{BE} = 0.1 and α_{LE} = 0.3. More details about the methodology used to tune image registration parameters can be found below.

The data term $\mathscr{L}$ is a linear combination of the local normalized cross correlation (LNCC)[43] and the squared euclidean distances between the landmarks positions

$$\mathscr{L}(I_{subject}, I_{atlas}, \phi(\Theta)) = \alpha_{LNCC} LNCC(I_{subject}, I_{atlas} \circ \phi(\Theta)) + \alpha_{LMKS} \sum_{k \in \Omega_{LMKS}} \left\| \phi(\Theta)(x_k^{subject}) - x_k^{atlas} \right\|^2 \tag{9}$$

where Ω_{LMKS} is the set of landmarks that are present for both $I_{subject}$ and I_{atlas}, $\alpha_{LMKS} = 0.001$ and $\alpha_{LNCC} = (1 - \alpha_{\text{LMKS}})(1 - \alpha_{BE} - \alpha_{LE})$ as implemented in `NiftyReg`[41]. The standard deviation of the Gaussian kernel of the LNCC was set to 6 mm. More details about the methodology used to tune image registration parameters can be found below.

Implementation details: Registrations that solve the optimization problem (7) were computed using the publicly available code for `NiftyReg`[41]. We used the latest version of the code on the master branch (git commit *99d584e*). The transformation ϕ in (7) is parameterized by cubic B-Splines of order 3 with a grid spacing equal to 3 mm. `NiftyReg`[41] uses a pyramidal approach to solve (7). We used 3 levels of pyramid which is the default value in `NiftyReg`. The brain mask were used to mask the voxels outside the brain.

The transformation ϕ in (7) was initialized with an affine transformation. The affine transformation was computed using a symmetric block-matching approach[44] based on image intensities and the brain masks. The implementation of the affine image registration method is included in `NiftyReg`.

Parameters tuning: The parameters α_{BE}, α_{LE}, α_{LMKS}, and the standard deviation of the Gaussian kernel of the LNCC of Equation (8) and Equation (9) were tuned using a grid search. The other parameters of the image registration were not tuned. The values of α_{BE} were {0.001, 0.01, 0.03, 0.1, 0.3}, the values of α_{LE} were {0.01, 0.03, 0.1}, the values of α_{LMKS} were {0.0003, 0.001, 0.003}, and the values for the standard deviation of the LNCC were {1, 2, 4, 6, 8}. We also tried to use the normalized mutual information (NMI) in place of the LNCC. There are no additional hyper-parameters related to NMI.

We selected the best set of parameter values using a subset of 22 pairs of 3D reconstructed MRIs covering the range of gestational ages available. The selection criteria was the mean of the Dice scores for the white matter, the ventricular system, and the cerebellum between volumes after non-linear registration. Details about the segmentation protocol can be found in section Semi-automatic segmentation of the atlas.

It is worth noting that the gradients of the different terms of the objective function in (7) have different scales. Therefore, comparing the contribution of the different terms based on their weights is misleading. Our parameter

tuning protocol suggests that all the terms of the objective function are important to obtain optimal image registration results. In particular, this supports the usefulness of the landmarks for the registration since a non-minimal value of α_{LMKS} was optimal.

3.3 Semi-automatic segmentation of the atlas

In this section, we describe the semi-automatic method that was used to obtain the segmentation for the proposed spatio-temporal atlas for SBA.

The fetal brains were divided into a total of eight tissue types: white matter (excluding the corpus callosum), ventricular system with the cavum septi pellucidi and cavum vergae, cerebellum, extra-axial cerebrospinal fluid, cortical grey matter, deep grey matter, brainstem, and corpus callosum. A visualization of the segmentations of those tissue types can be found in Figure 11 and Figure 12. The annotation protocol follows the annotation guidelines of the FeTA dataset[23]. In addition, the corpus callosum was also delineated.

Automatic 3D tissue types probability maps were obtained using a deep learning pipeline trained using partially supervised learning[26]. An ensemble of ten deep neural networks trained using the Leaf-Dice loss[26] has been used. The code and the pre-trained networks used for the automatic segmentation are available here. An average 3D tissue types probability maps for the atlas was obtained using a weighted average method analogous to the one described in section Time-weighted average of the volumes for the 3D reconstructed MRIs. Formally, let $\{P_i\}_{i=1}^{N}$ be a set of N co-registered 3D tissue types probability maps to average. The weighted average is computed as

$$P_{average} = \frac{1}{2N}\sum_{i=1}^{N} w_i (P_i + S(P_i)) \tag{10}$$

where S is the operator that computes the symmetric of a volume with respect to the central sagittal and the weights w_i are defined as in section Time-weighted average of the volumes. An initial segmentation of the atlas was obtained using the tissue types of maximum probability for each voxel.

The initial segmentations of the spatio-temporal atlas were quality controlled and corrected when necessary by authors LF and MA, a paediatric radiologist specialized in fetal brain anatomy with eight years of experience in segmenting fetal brain MRIs. Manual segmentations were performed using the software ITK-SNAP[34] version 3.8.0.

4 Annotation potocol of anatomical landmarks for fetuses with spina bifida aperta

In this section, protocols designed for the selection of imaging landmarks in MRI images of fetal brains with spina bifida aperta (SBA) are outlined. This is aimed to improve the accuracy of image registration. A total of 11 anatomical landmarks per study have been selected for initial assessment. Four in each cerebral hemisphere and three in the posterior fossa.

The first seven landmarks described below were found to be sufficiently reliable. The last four landmarks involving the cavum septi pellucidi were found to be insufficiently reliable.

4.1 Anterior horn of the right lateral ventricle

In the axial plane identify the right lateral ventricle. Use the view in the sagittal plane to select the most anterior slice reached by the ventricle. When this slice is not unique, which occurs when the anterior border of the ventricle is flattened, select the slice at the centre. The border is considered as the brighter intensity value of the two lines of intensity values showing the greatest difference. An illustration is given in Figure 4.

4.2 Anterior horn of the left lateral ventricle

In the axial plane identify the Left Lateral Ventricle. Use the view in the sagittal plane to select the most anterior slice reached by the ventricle. When this slice is not unique, which occurs when the anterior border of the ventricle is flattened, select the slice at the centre. The border is considered as the brighter intensity value of the two lines of intensity values showing the greatest difference. An illustration is given in Figure 4.

4.3 Posterior tectum plate

Using the sagittal and axial planes locate the tectum. In the axial plane select the midline sagittal slice. Confirm using the sagittal plane that the axial slice is viewing the most prominent part of the tectum. Using the smallest marker select the most posterior point of the tectum tissue. This considered to be the lower intensity value of the two intensity values at the posterior peak showing the greatest difference. An illustration is given in Figure 5.

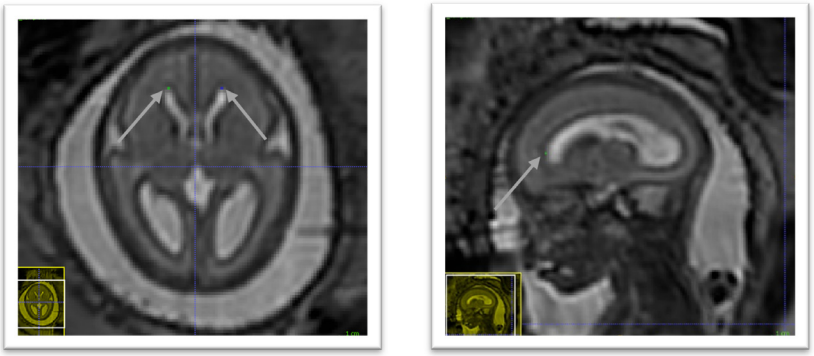


Figure 4. Anterior horn of the right lateral ventricle (green) and anterior horn of the left lateral ventricle (blue).

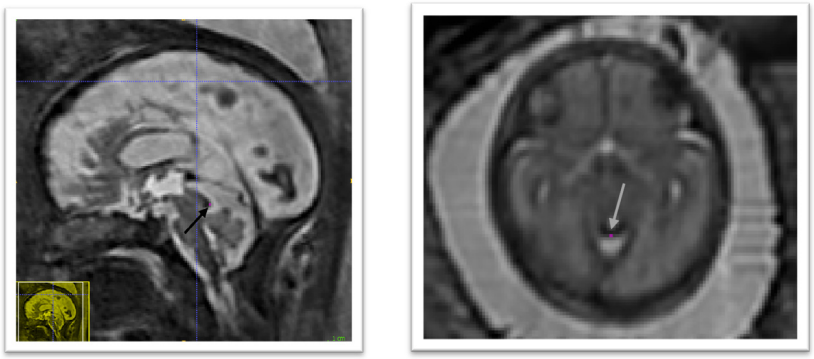


Figure 5. Posterior tectum plate (pink).

4.4 Left cerebellar-brainstem junction

In the axial view we locate the cerebellum and select the slice with the greatest cerebellar width, preferably where the posterior fossa also is seen at its greatest width. The brainstem is found just anterior to the cerebellum and directly meets with the cerebellum along its posterior borders. In this area, we select with the smallest possible marker the most anterior point where the cerebellum and brainstem meet on the left side. The marker should be within cerebellar tissue as oppose to the tissue of the brainstem. An illustration is given in Figure 6.

4.5 Right cerebellar-brainstem junction

In the axial view we locate the cerebellum and select the slice with the greatest cerebellar width, preferably where the posterior fossa also is seen at its greatest width. The brainstem is found just anterior to the cerebellum and directly meets with the cerebellum along its posterior borders. In this area, we select with the smallest possible marker the most anterior point where the cerebellum and brainstem meet on the right side. The marker should be within cerebellar tissue as oppose to the tissue of the brainstem. An illustration is given in Figure 6.

4.6 Left deep grey border at foramen of Monro

In the axial view locate the foramen of Monro or the interventricular foramen. The paired foramina connect the lateral ventricles to the third ventricle. The point where the foramina lead into the third ventricle, a horseshoe or trough shaped border is formed anteriorly. If not visible in this way, it can also be observed in the coronal view connecting the anterior horns of the lateral ventricle to the third ventricle. Select the mid-sagittal slice and trace a horizontal line left across from this border. The correct position of the line is considered as the row of brighter intensity value of the two rows of intensity values showing the greatest contrast. The edge of the deep grey matter on the left side which should be visible forming a darker grey arch from the left anterior horn to the left posterior horn of the lateral ventricles. Using the smallest possible marker, mark the edge of the deep grey matter where it intersects with the line. An illustration is given in Figure 7.



Figure 6. Right cerebellar-brainstem junction (turquoise) and left cerebellar-brainstem junction (pink).

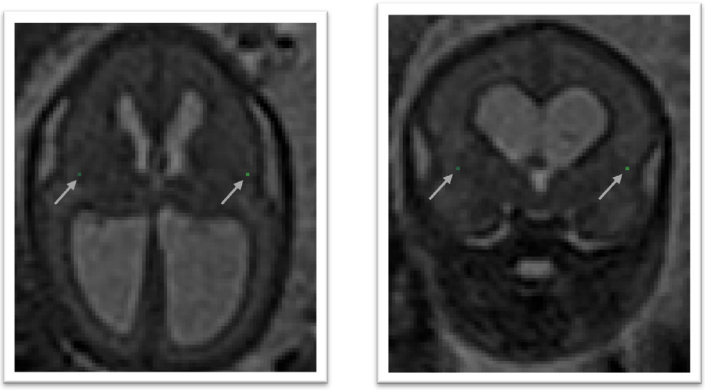


Figure 7. Left deep grey matter border at foramen of Monro (dark olive) and right deep grey matter border at foramen of Monro (lime green).

4.7 Right deep grey border at foramen of Monro

In the axial view locate the foramen of Monro or the interventricular foramen. The paired foramina connect the lateral ventricles to the third ventricle. The point where the foramina lead into the third ventricle, a horseshoe or trough shaped border is formed anteriorly. If not visible in this way, it can also be observed in the coronal view connecting the anterior horns of the lateral ventricle to the third ventricle. Select the mid-sagittal slice and trace a horizontal line right across from this border. The correct position of the line is considered as the row of brighter intensity value of the two rows of intensity values showing the greatest contrast. The edge of the deep grey matter on the right side which should be visible forming a darker grey arch from the right anterior horn to the right posterior horn of the lateral ventricles. Using the smallest possible marker, mark the edge of the deep grey matter where it intersects with the line. An illustration is given in Figure 7.

4.8 Left deep grey border at anterior cavum septi pellucidi line

In the axial view locate the cavum septi pellucidi (CSP), a cavity in the fetal brain, the leaflets of the septum pellucidum are located between the anterior horns of the lateral ventricles. Select the slice in which the anterior wall of the cavity is found most anteriorly. If there is significant abnormality in this structure it may be helpful to use the sagittal plane to assist in defining this area. Trace a horizontal line left across from the anterior wall of the cavum septi pellucidi. The correct position of the line is considered as the row of brighter intensity value of the two rows of intensity values showing the greatest contrast. The edge of the deep grey matter on the left side forms a darker arch from the left anterior horn to the left posterior horn of the lateral ventricles. Using the smallest possible marker mark the edge of the deep grey matter where it intersects with that line. An illustration is given in Figure 8.

4.9 Right deep grey border at the anterior cavum septi pellucidi line

In the axial view locate the cavum septi pellucidi (CSP), a cavity in the fetal brain, the leaflets of the septum pellucidum are located between the anterior horns of the lateral ventricles. Select the slice in which the anterior

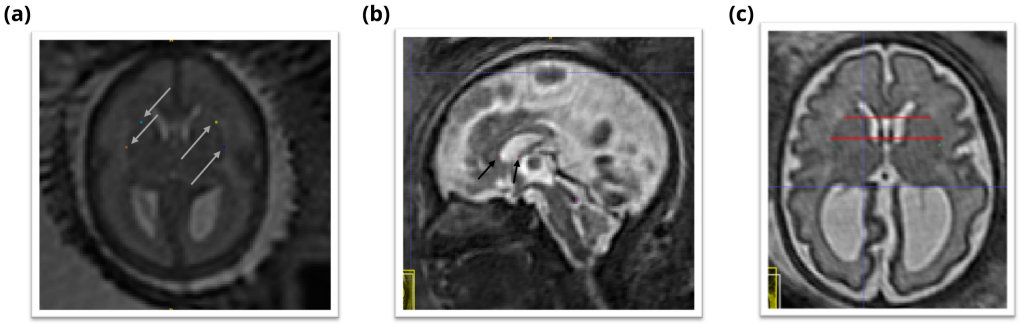


Figure 8. Deep grey matter border with respect to the cavum septi pellucidi (CSP). (**a**) Left deep grey border at anterior CSP line (yellow), right deep grey border at the anterior CSP line (light blue), left deep grey border at posterior CSP line (dark blue), right deep grey border at the posterior CSP line (orange). (**b**) Sagittal view of the position of the horizontal lines used to guide the marking of the deep grey borders at CSP (red). (**c**) Axial view of the position of the horizontal lines used to guide the marking of the deep grey borders at CSP (red).

wall of the cavity is found most anteriorly. If there is significant abnormality in this structure it may be helpful to use the sagittal plane to assist in defining this area. Trace a horizontal line right across from the anterior wall of the cavum septi pellucidi. The correct position of the line is considered as the row of brighter intensity value of the two rows of intensity values showing the greatest contrast. The edge of the deep grey matter on the right side forms a darker arch from the right anterior horn to the right posterior horn of the lateral ventricles. Using the smallest possible marker mark the edge of the deep grey matter where it intersects with that line. An illustration is given in Figure 8.

4.10 Left deep grey border at posterior cavum septi pellucidi line

In the axial view locate the cavum septi pellucidi, a cavity in the fetal brain, the leaflets of the septum pellucidum are located between the anterior horns of the lateral ventricles. Select the slice in which the anterior wall of the cavity is found most anteriorly. If there is significant abnormality in this structure it may be helpful to use the sagittal plane to assist in defining this area. At this level trace a horizontal line left across from the posterior wall of the cavum septi pellucidi. The correct position of the line is considered as the row of brighter intensity value of the two rows of intensity values showing the greatest contrast. The edge of the deep grey matter on the left side forms a darker arch from the left anterior horn to the left posterior horn of the lateral ventricles. Using the smallest possible marker mark the edge of the deep grey matter where it intersects with that line. An illustration is given in Figure 8.

4.11 Right deep grey border at the posterior cavum septi pellucidi line

In the axial view locate the cavum septi pellucidi, a cavity in the fetal brain, the leaflets of the septum pellucidum are located between the anterior horns of the lateral ventricles. Select the slice in which the anterior wall of the cavity is found most anteriorly. In this slice trace a horizontal line right across from the posterior wall of the cavum septi pellucidi. The correct position of the line is considered as the row of brighter intensity value of the two rows of intensity values showing the greatest contrast. The edge of the deep grey matter on the right side forms a darker arch from the right anterior horn to the right posterior horn of the lateral ventricles. Using the smallest possible marker mark the edge of the deep grey matter where it intersects with that line. An illustration is given in Figure 8.

5 Results

5.1 Intra-rater variability for the annotation of the anatomical landmarks

To assess intra-rater variability, a subset of 31 3D reconstructed MRIs, selected at random, were marked two times by the same rater, EV. The mean gestational age was 26.2 weeks and the range of gestational ages in the reliability set was 22–34 weeks. Those statistics closely match the one of the full cohort as described in section Spina bifida aperta cohort used to compute the spatio-temporal atlas (the mean gestational age is 26.1 weeks and the range is 21 – 35 weeks for the full dataset). The two ratings were performed with an interval of at

least three weeks to mitigate the bias caused by observer recollection. A landmark was marked absent when the anatomical position described by the protocol was not found within the volume.

The two landmark placements are said to be in agreement if the second landmark placement is inside a 3 × 3 × 3 voxel cube where the original placement is the central voxel. When 95% of the second landmarks fall within this radius, the landmark is considered 'Excellent' in terms of intra-rater reliability, when 80% of are in agreement, intra-rater reliability is considered 'Good', where 75% fall within the radius of agreement intra-rater reliability is considered 'Satisfactory'. For landmarks with a probability of agreement of less than 75%, the reliability is considered 'Poor'. The probabilities that pairs of landmarks are in agreement is estimated based on the assumption that the distribution of distances between first and second marks is Gaussian. The results can be found in Table 1.

5.2 Automatic segmentation of fetal brain 3D MRIs

In this section, we compare the automatic segmentations obtained either using an atlas of normal fetal brains[20] or using the proposed atlas for spina bifida aperta (SBA). The quantitative evaluation can be found in Table 2.

We studied the automatic segmentation of fetal brain 3D MRIs into seven tissue types and brain extraction[45]. Fetal brain 3D MRIs from the FeTA dataset[23,25] were used for the evaluation. More details about the dataset used for the evaluation can be found in section Fetal brain 3D MRI used for the evaluation of automatic segmentation.

The automatic segmentations are obtained in two steps: first a volume of the atlas, chosen based on the gestational age, is registered to each fetal brain 3D MRI, and second, after registration, the segmentation of the atlas is propagated. Non-linear image registration is implemented as described in section Non-linear image registration. In particular, we used the same hyper-parameter values and the anatomical landmarks are not used

Table 1. Evaluation of the reliability of the landmarks. We report the estimated percentiles of distances in millimeters between first and second marking for each proposed landmarks. P_{75}: 75th percentile of distances in millimeters. P_{80}: 80th percentile of distances in millimeters. P_{95}: 95th percentile of distances in millimeters. Our reliability score is defined in section 5.1. **LALV**: Anterior Horn of the Left Lateral Ventricle, **RALV**: Anterior Horn of the Right Lateral Ventricle, **PTP**: Posterior Tectum Plate, **LCB**: Left Cerebellar Brainstem Junction, **RCB**: Right Cerebellar Brainstem Junction, **LFOM**: Left Deep Grey Border at Foramen of Monro, **RFOM**: Right Deep Grey Border at Foramen of Monro, **LACSP**: Left Deep Grey Border at Anterior Cavum Septi Pellucidi line, **RACSP**: Right Deep Grey Border at Anterior Cavum Septi Pellucidi line, **LPCSP**: Left Deep Grey Border at Posterior Cavum Septi Pellucidi line, **RPCSP**: Right Deep Grey Border at the Posterior Cavum Septi Pellucidi line.

Landmark	Ratio of Missing (%)	P_{75} (mm)	P_{80} (mm)	P_{95} (mm)	Reliability
LALV	0	1.73	1.95	3.02	Good
RALV	0	1.70	1.91	2.96	Good
PTP	3	1.15	1.29	2.00	Excellent
LCB	0	1.70	1.90	2.95	Good
RCB	0	1.78	2.00	3.10	Good
LFOM	3	2.83	3.17	4.91	Poor
RFOM	0	2.50	2.81	4.35	Satisfactory
LACSP	16	2.74	3.07	4.77	Poor
RACSP	29	2.59	2.91	4.51	Satisfactory
LPCSP	16	3.35	3.76	5.83	Poor
RPCSP	16	3.12	3.50	5.43	Poor

Table 2. Evaluation of automatic fetal brain segmentation. We report mean (standard deviation) for the Dice score (DSC) in percentages and the Hausdorff distance at 95% (HD95) in millimeters for all tissue types. **Brain:** whole brain that includes all the tissue types below, **WM:** white matter, **Vent:** ventricular system, **Cer:** cerebellum, **CSF:** cerebrospinal fluid, **ECSF:** extra-axial CSF, **CGM:** cortical grey matter, **DGM:** deep grey matter, **BS:** brainstem.

Atlas	Cohort	Metric	Brain	WM	Vent	Cer	ECSF	CGM	DGM	BS
Normal[20]	Normal	DSC	97.5 (1.2)	89.5 (3.2)	84.2 (3.5)	89.2 (3.8)	87.8 (3.6)	74.4 (8.0)	85.2 (3.7)	82.0 (3.4)
		HD95	1.5 (0.9)	1.7 (0.9)	1.3 (0.6)	1.4 (0.4)	1.3 (0.7)	1.4 (0.9)	2.2 (0.7)	2.3 (0.5)
Normal[20]	Spina Bifida	DSC	91.4 (11.4)	69.4 (19.5)	76.6 (17.3)	53.7 (32.2)	52.0 (36.5)	45.0 (25.0)	69.9 (19.7)	62.1 (23.9)
		HD95	3.6 (3.4)	4.5 (3.4)	4.0 (3.8)	7.0 (7.3)	10.7 (10.6)	4.2 (3.5)	4.3 (3.4)	4.8 (5.5)
Spina Bifida	Spina Bifida	DSC	92.8 (4.9)	83.3 (6.8)	87.1 (8.9)	74.3 (15.1)	59.9 (26.8)	54.6 (16.7)	79.2 (5.4)	73.0 (9.9)
		HD95	3.1 (1.7)	3.0 (1.5)	1.9 (1.3)	3.2 (5.2)	8.6 (9.1)	2.8 (1.5)	2.9 (0.8)	2.8 (1.0)

during the registration. The automatic segmentations for the corpus callosum and the white matter were merged into white matter, since the corpus callosum is part of the white matter segmentation in the FeTA dataset.

Automatic segmentations for the SBA cases are computed using either a normal fetal brain atlas[20] or our SBA fetal brain atlas as can be seen in the last four rows of Table 2. Segmentation results per gestational age for SBA cases can be found in Figure 9 and Figure 10. In addition, we have also computed automatic segmentations for the normal brain cases using the normal fetal brain atlas[20] as can be seen in the first two rows of Table 2. The evaluation was performed for each tissue type using the Dice score[46,47] and the Hausdorff distance at percentile 95[48].

6 Discussion

The proposed spatio-temporal atlas for spina bifida aperta (SBA) is illustrated in Figure 11 and Figure 12 (see *Data availability*[49] and *Software availability* for full atlas).

As described in section Spina bifida aperta cohort used to compute the spatio-temporal atlas, the cohort used to compute this atlas contains longitudinal data. This longitudinal dataset of 90 MRIs might be less representation of the whole SBA population than a dataset of 90 MRIs that would contain only singletons. However, the use of longitudinal data adds some implicit temporal consistency in the atlas.

The landmarks in the ventricles, the posterior tectum plate, and at the junction of the cerebellar and the brainstem were all found to be reliable enough in terms of distance between successive marks by the same rater as can be seen in Table 1. In addition, those anatomical landmarks were always present, except for the posterior tectum plate that was missing for one reconstructed 3D MRI. However, the landmarks in the deep grey were almost all found to be poorly reliable in terms of distance between successive marks by the same rater. One can group the landmarks in the deep grey matter into two groups: the landmarks based on the foramen of Monro, and the landmarks based on the cavum septi pellucidi. The landmarks based on the foramen of Monro were almost always present. This is in contrast with the landmarks based on the cavum septi pellucidi that were missing up to 29% of the time. In Figure 13, we give an illustration of the anatomical variability of the cavum septi pellucidi in fetuses with SBA. This suggests that the position of landmarks based on the cavum septi pellucidi can vary widely from one subject to the other. As a result, we choose to use the two landmarks based on the foramen of Monro for the computation of the atlas, but to exclude the four landmarks based on the cavum sceptum pellucidum.

The evaluation of automatic segmentation of fetal brain 3D MRIs in Table 2 suggests that using the proposed atlas for SBA leads to more accurate segmentation of SBA cases than a normal fetal brain atlas. The

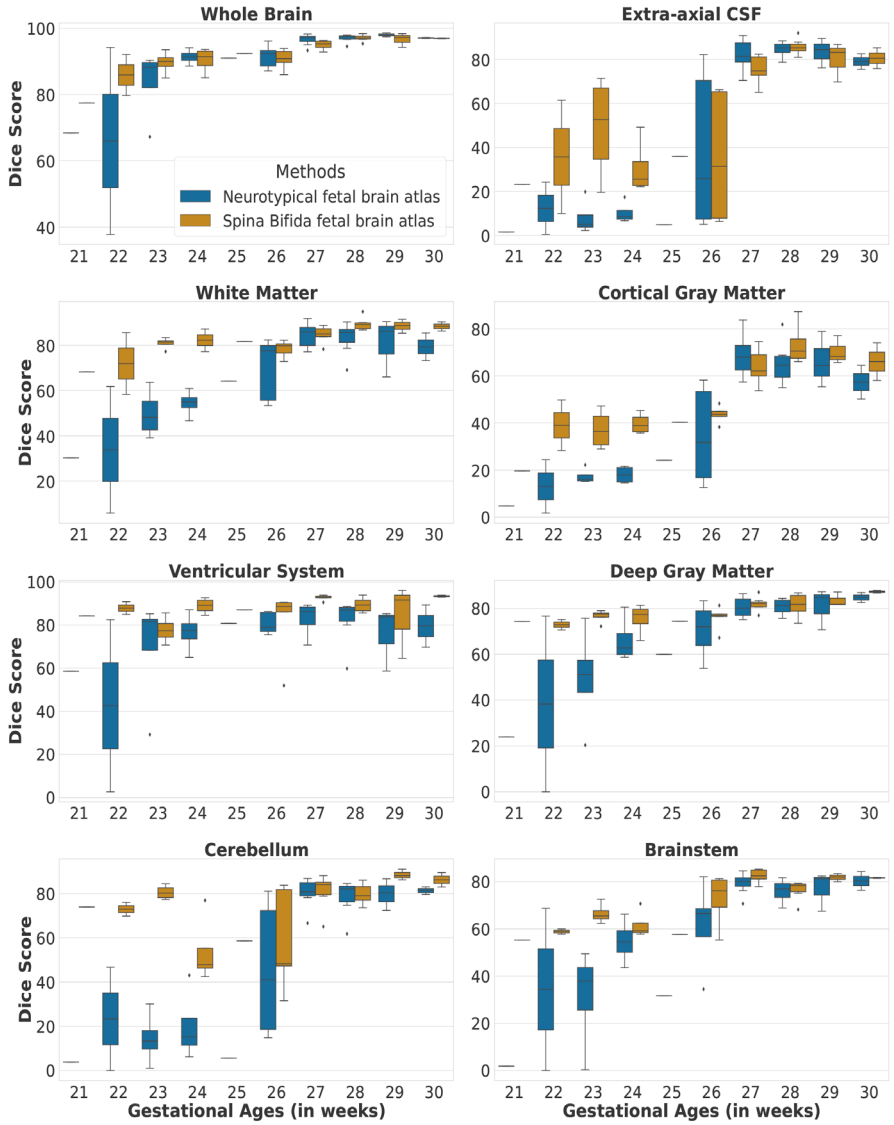


Figure 9. Dice scores per tissue type and per gestational age for the spinal bifida evaluation cohort (36 3D MRIs).

proposed atlas for SBA outperforms the normal fetal brain atlas in terms of mean Dice scores and mean Hausdorff distances for all tissue types. The proposed atlas also leads to lower standard deviations of Dice scores and Hausdorff distances for all tissue types. This suggests that automatic segmentation using image registration of an atlas is more robust for SBA when an SBA atlas is used. We have investigated the segmentation performance for SBA per tissue type and per gestational age in Figure 9 and Figure 10. We can observe that the proposed SBA atlas outperforms the normal fetal brain atlas by the largest margins for gestational ages 25 weeks or lower. The week 27 is the only week for which the proposed spina bifida atlas underperforms the baseline for the extra-axial CSF and the deep grey matter. An artefact visible in the orbito-frontal region in Figure 12 may account for the suboptimal automatic segmentation at 27 weeks.

With fetal surgery the open neural tube defect is closed and thus the continuous leakage of CSF is stopped. This leads to an increase in CSF within the skull, surrounding the cerebrum and cerebellum, leading to a better distinction of the grey matter from the inner lining of the skull. This may explain the higher segmentation accuracy for the grey matter after 27 weeks as from this time point the evaluation cohort includes only post-operative MRIs. In addition to the increase in CSF within the skull, the closure of the defect leads to a reversal

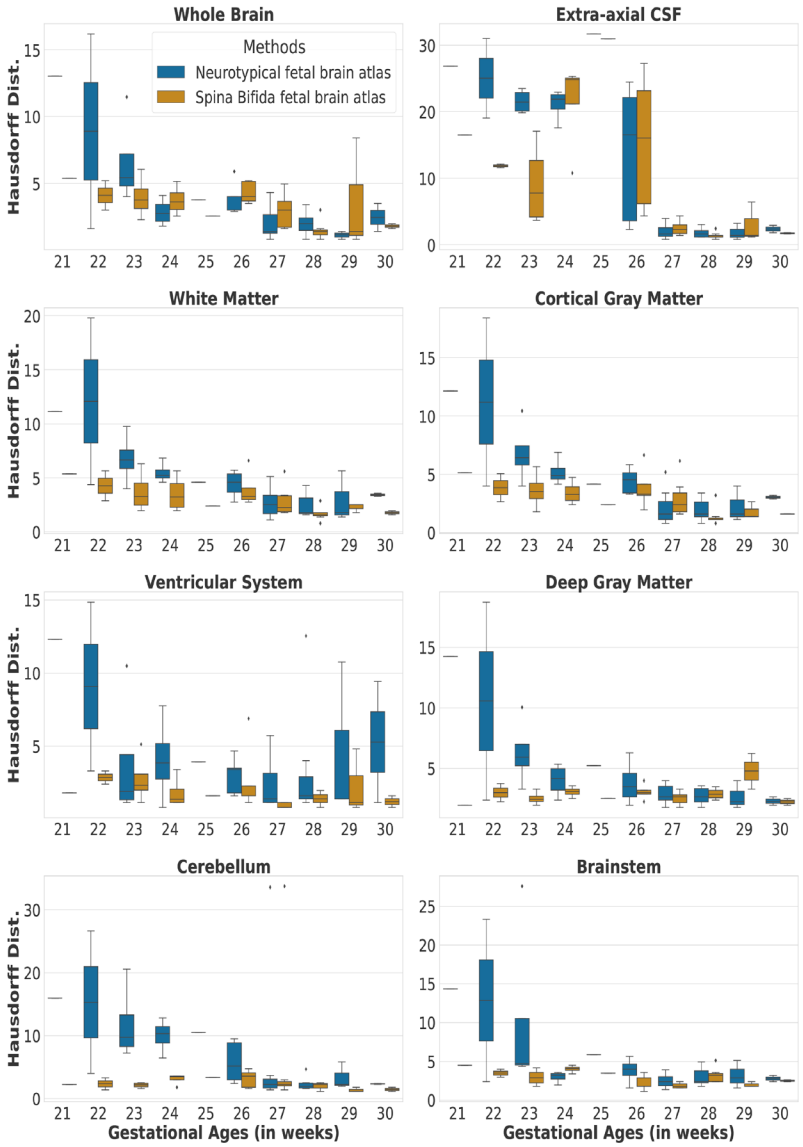


Figure 10. Hausdorff distances per tissue type and per gestational age for the spinal bifida evaluation cohort (36 3D MRIs).

of the hindbrain herniation. This happens already within 7 days after surgery in the majority of cases[31]. The reversal of the hindbrain information in combination with the increase in fluid surrounding the cerebellum and brainstem in the posterior fossa improves the distinction of the cerebellum and brainstem from the skull base. We notice an improved performance after 27 weeks, supporting the impact of fluid restoration in the skull on our automatic segmentation algorithm.

In addition, when comparing automatic segmentations of normal fetuses and fetuses with SBA obtained using a normal fetal brain atlas we found a decrease of segmentation accuracy in terms of Dice scores and Hausdorff distances for all tissue types. For the cerebellum, the mean Dice score decreased from 89.2% for normal fetuses to 53.7% for fetuses with SBA. This can be attributed to the Chiari malformation type II which is found in most SBA cases[2]. The decrease of mean Dice score and the increase of mean Hausdorff distance for the extra-axial cerebrospinal fluid (CSF) can be attributed to the quasi absence of extra-axial CSF in fetuses with SBA at early developmental stages as illustrated in Figure 9 and Figure 10.

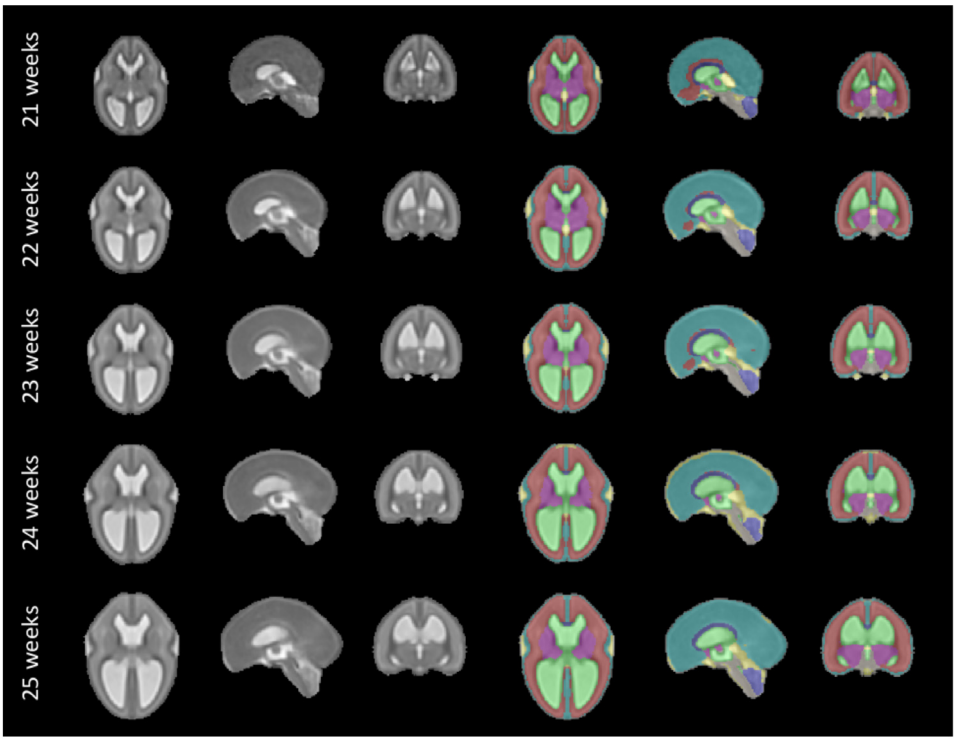


Figure 11. Our spatio-temporal atlas for spina bifida aperta - Part I (not operated). Publicly available here.

It is worth noting the large variability of the segmentation results for week 26 in Figure 9 and Figure 10 for the two atlases. This can be attributed to the variability in the topology of the extra-axial CSF illustrated in Figure 14. At week 26, the spina bifida atlas performs best on 3D MRIs of fetuses with limited extra-axial CSF (Figure 14 left) while the neurotypical atlas performs best for 3D MRIs of fetuses with circumferential extra-axial CSF (Figure 14 right)

7 Limitations

In this work, we have used MRIs of operated and non-operated fetuses, ie that have or have not undergone fetal surgery to close the spina bifida aperta (SBA) defect in utero. *In-utero* fetal surgery is currently recommended to be performed prior to 26 weeks of gestation. The surgery has been found to influence the evolution of the fetal brain anatomy starting within one week after the operation[31]. Therefore, a normative atlas for SBA should be computed using only MRIs of non-operated fetuses. This limitation of our work is however due to the clinical data used. To make this limitation clear we have separated the atlas into two parts as illustrated in Figure 11 and Figure 12. This separations is also reflected in the data structure chosen to share the atlas, as detailed in Underlying data[49].

In Figure 1, it is worth noting that relatively little cases are available in the range of gestational ages 27 – 31 weeks. As a result, the proposed atlas might be less representative of the SBA population in this range of gestational ages. In particular, this might explain why the ventricle size does not appear to increase linearly for those gestational ages as can be seen in Figure 12.

An artefact is visible on the 27-week atlas in the orbito-frontal region. The artefact in the orbito-frontal region may account for the suboptimal automatic segmentation at 27 weeks. We think that this artefact is due to variability in the topology of the extra-axial cerebrospinal fluid (CSF) at 26 weeks for operated fetuses. At this time point, the extra-axial CSF can be either circumferential or limited as illustrated in Figure 14 that we have added. In this case, the diffeomorphic hypothesis of our non-linear registration step is violated and can lead to such artefacts. It is worth noting that this issue affects the atlas at 27 weeks due to the use of a time-weighted regression kernel.

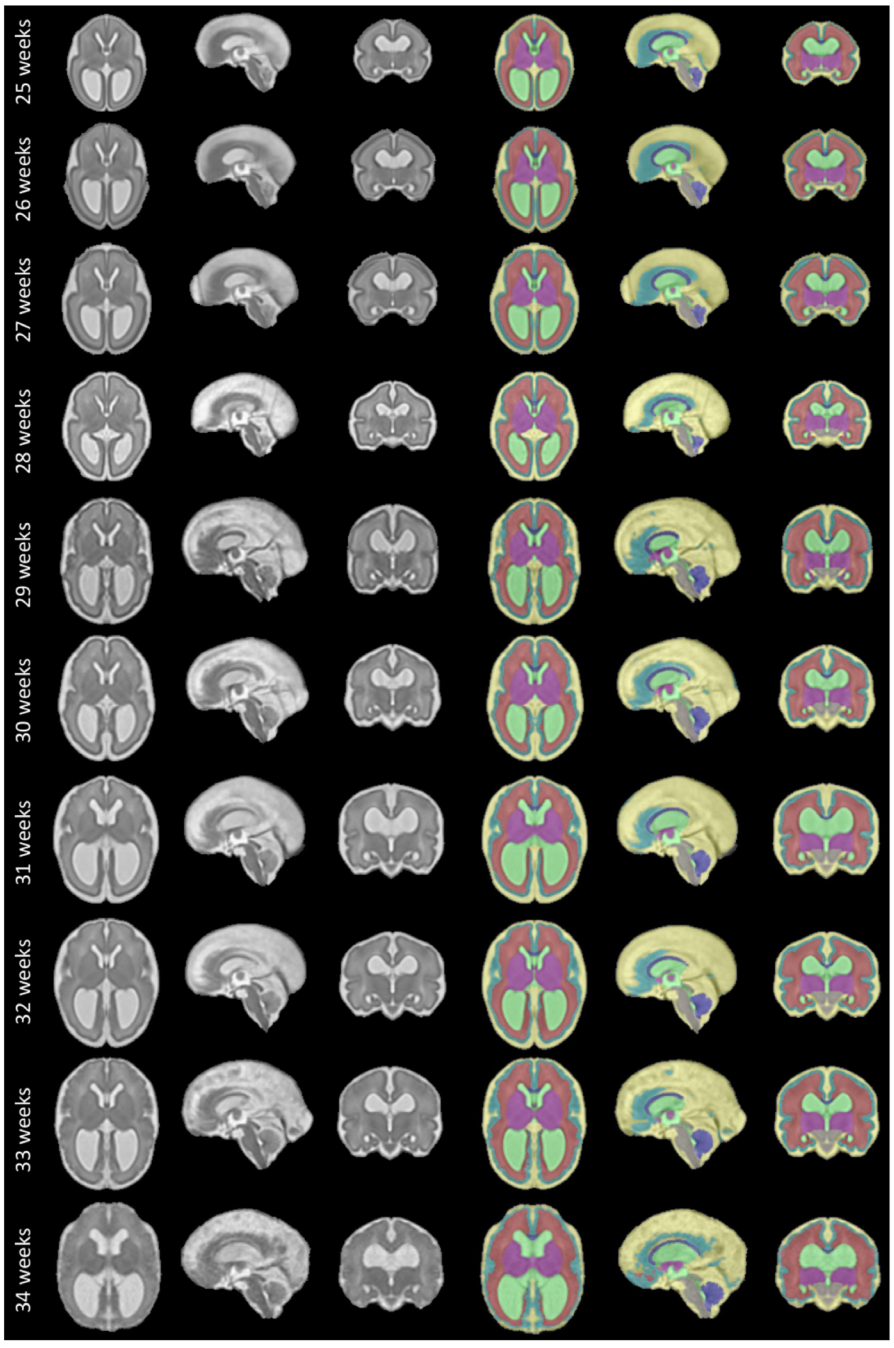


Figure 12. Spatio-temporal atlas for spina bifida aperta - Part II (operated). Publicly available here.

The proposed atlas mixes male and female fetuses. However, recent work reported different brain growth trajectories between male and female neurotypical fetuses[50]. Sex-specific atlases may be of interest to better represent the populations of male and female fetuses with spina bifida aperta.

Regarding the evaluation, the spina bifida 3D MRIs of the FeTA dataset cover only the gestational ages from 20 weeks to 30 weeks. As a result, the segmentation accuracy obtained using the atlases for gestational ages higher than 30 weeks was not evaluated.

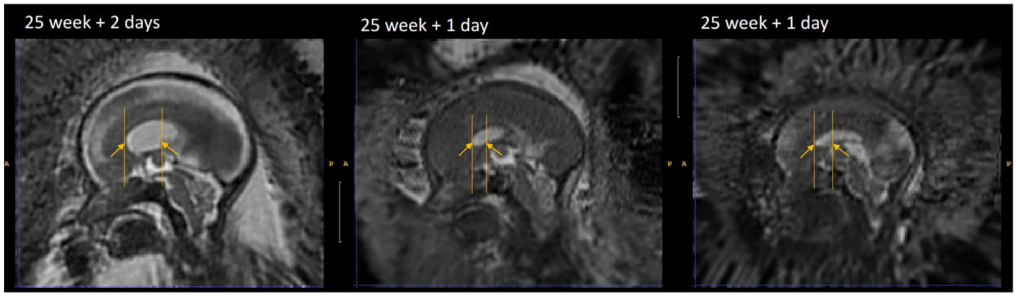


Figure 13. Cavum septi pellucidi (CSP) variation fetuses with 25 weeks of gestation. Yellow arrows indicate the anterior and posterior borders of the CSP as defined by the landmark localisation protocol. This visualisation illustrates the disparity between volumes in terms of shape and size of the CSP.

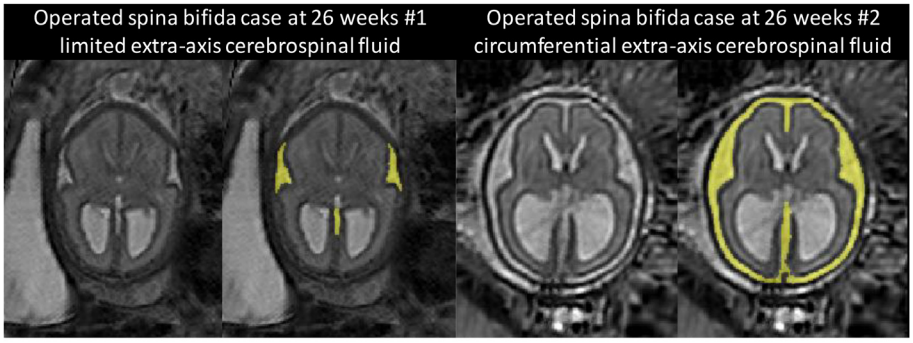


Figure 14. Two 3D MRIs of operated SBA fetuses at 26 weeks with different extra-axial CSF topologies. The extra-axial CSF is highlighted in yellow. This variation of topology has an impact on the segmentation accuracy and may be the cause of an artefact in the SBA atlas at week 27.

8 Conclusions

In this work we propose the first spatio-temporal fetal brain MRI atlas for spina bifida aperta (SBA).

We propose a semi-automatic pipeline for the computation of spatio-temporal fetal brain atlas. Our pipeline relies on four main components:

- MONAIfbs[29], an automatic method for fetal brain extraction in 2D fetal MRIs.
- NiftyMIC[30], a 3D super resolution and reconstruction algorithm that allows to obtain isotropic and motion-free volumetric MRI of the fetal brain.
- A proposed protocol for the annotation of 7 anatomical landmarks in 3D reconstructed fetal brain MRIs.
- A proposed weighted generalize Procrustes method for an unbiased initialization of the atlas based on the anatomical landmarks.

We find that the proposed atlas outperforms a state-of-the-art fetal brain atlas for the automatic segmentation of brain 3D MRIs of fetuses with SBA. This suggests that the proposed atlas for SBA provides a better anatomical prior about the peri-surgical SBA brain. We hypothesise that this atlas could also help improving fetal brain MRI segmentation methods that lacks such prior, such as segmentation methods based on deep learning[26]. We are planning to investigate this in the future.

Data availability

Underlying data

Ethical approval allows us to use the magnetic resonance imaging (MRI) data from University Hospitals Leuven for research and to make publicly available results obtained using those data such as the fetal brain atlas

for SBA proposed in this work. The Caldicott guardian at University College London Hospital (UCLH) gave their approval to share the data with University College London and King's College London researchers for analysis. However, we do not have the required ethical approval to share the original MRI data publicly. Readers and reviewers can email the corresponding author (lucas.fidon@kcl.ac.uk) to request access to the data. Access to the data at UCLH will require approval by the Caldicott guardian at UCLH and access to the data from University Hospitals Leuven will require approval by the ethics committee at University Hospitals Leuven.

The FeTA dataset is publicly available on Synapse: https://doi.org/10.7303/syn23747212. Access requires registration to Synapse and agreement to the terms of use.

The manual segmentations for the fetal brain MRI of FeTA dataset, that we have contributed in our previous work[26–28], are publicly available on Zenodo: https://doi.org/10.5281/zenodo.6878474[51] under the term of the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported license (CC BY-NC-ND 3.0). Access to the data is restricted. Readers and reviewers can apply for access to the data by filling in a form. The only requirement is to acknowledge that the applicant will not use those data for commercial purposes.

The spatio-temporal atlas of the normal developing fetal brain that we have used for comparison is publicly available at http://crl.med.harvard.edu/research/fetal_brain_atlas/. Access requires readers to fill in an access form. Alternatively, one can download the fetal brain atlas directly from the `NiftyMIC` GitHub repository.

Zenodo: A Spatio-temporal Atlas of the Developing Fetal Brain with Spina Bifida Aperta. https://doi.org/10. 5281/zenodo.5524312[49].

This project contains the following underlying data:

The project contains 15 folders, each corresponding to a unique volume of our spatio-temporal fetal brain atlas, as illustrated in Figure 11 and Figure 12, and contains four nifti files:

- srr.nii.gz (average 3D reconstructed MRI).
- mask.nii.gz (3D brain mask).
- parcellation.nii.gz (3D segmentation of the fetal brain into 8 tissue types as described in section Semi-automatic segmentation of the atlas).
- lmks.nii.gz (annotations for the 7 anatomical landmarks described is section Anatomical landmarks).

Data are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication). Codes and scripts are available under the terms of the BSD-3-Clause license.

Alternatively, it is possible to download A Spatio-temporal Atlas of the Developing Fetal Brain with Spina Bifida Aperta on Synapse: https://doi.org/10.7303/syn25887675. It is necessary to create a synapse account to be able to download the data.

Software availability

Source code available from: https://github.com/LucasFidon/spina-bifida-MRI-atlas

Archived source code at the time of publication: https://doi.org/10.5281/zenodo.5524312[49]

License: BSD-3-Clause

Acknowledgements

We would like to thank Bella Spencer for valuable feedback on an early version of the plain language summary.

References

1. Khoshnood B, Loane M, de Walle H, *et al.*: **Long term trends in prevalence of neural tube defects in Europe: population based study.** *BMJ.* 2015; **351**: h5949.
PubMed Abstract | Publisher Full Text | Free Full Text
2. Pollenus J, Lagae L, Aertsen M, *et al.*: **The impact of cerebral anomalies on cognitive outcome in patients with spina bifida: A systematic review.** *Eur J Paediatr Neurol.* 2020; **28**: 16–28.
PubMed Abstract | Publisher Full Text
3. Naidich TP, Pudlowski RM, Naidich JB, *et al.*: **Computed tomographic signs of the Chiari II malformation. Part I: Skull and dural partitions.** *Radiology.* 1980; **134**(1): 65–71.
PubMed Abstract | Publisher Full Text
4. Kunpalin Y, Deprest J, Papastefanou I, *et al.*: **Incidence and patterns of abnormal corpus callosum in fetuses with isolated spina bifida aperta.** *Prenat Diagn.* 2021; **41**(8): 957–964.
PubMed Abstract | Publisher Full Text
5. Crawley JT, Hasan K, Hannay HJ, *et al.*: **Structure, integrity, and function of the hypoplastic corpus callosum in spina bifida myelomeningocele.** *Brain Connect.* 2014; **4**(8): 608–618.
PubMed Abstract | Publisher Full Text | Free Full Text
6. Dennis M, Cirino PT, Simic N, *et al.*: **White and grey matter relations to simple, choice, and cognitive reaction time in spina bifida.** *Brain Imaging Behav.* 2016; **10**(1): 238–251.
PubMed Abstract | Publisher Full Text
7. Treble-Barna A, Juranek J, Stuebing KK, *et al.*: **Prospective and episodic memory in relation to hippocampal volume in adults with spina bifida myelomeningocele.** *Neuropsychology.* 2015; **29**(1): 92–101.
PubMed Abstract | Publisher Full Text | Free Full Text
8. Mufti N, Aertsen M, Ebner M, *et al.*: **Cortical spectral matching and shape and volume analysis of the fetal brain pre- and post-fetal surgery for spina bifida: a retrospective study.** *Neuroradiology.* 2021; **63**(10): 1721–1734.
PubMed Abstract | Publisher Full Text | Free Full Text
9. Treble A, Juranek J, Stuebing KK, *et al.*: **Functional significance of atypical cortical organization in spina bifida myelomeningocele: relations of cortical thickness and gyrification with IQ and fine motor dexterity.** *Cereb Cortex.* 2013; **23**(10): 2357–2369.
PubMed Abstract | Publisher Full Text | Free Full Text
10. Hasan KM, Sankar A, Halphen C, *et al.*: **Quantitative diffusion tensor imaging and intellectual outcomes in spina bifida: laboratory investigation.** *J Neurosurg Pediatr.* 2008; **2**(1): 75–82.
PubMed Abstract | Publisher Full Text | Free Full Text
11. Mandell JG, Kulkarni AV, Warf BC, *et al.*: **Volumetric brain analysis in neurosurgery: Part 2. Brain and CSF volumes discriminate neurocognitive outcomes in hydrocephalus.** *J Neurosurg Pediatr.* 2015; **15**(2): 125–132.
PubMed Abstract | Publisher Full Text
12. Mufti N, Sacco A, Aertsen M, *et al.*: **What brain abnormalities can magnetic resonance imaging detect in foetal and early neonatal spina bifida: a systematic review.** *Neuroradiology.* 2022; **64**(2): 233–245.
PubMed Abstract | Publisher Full Text | Free Full Text
13. Danzer E, Joyeux L, Flake AW, *et al.*: **Fetal surgical intervention for myelomeningocele: lessons learned, outcomes, and future implications.** *Dev Med Child Neurol.* 2020; **62**(4): 417–425.
PubMed Abstract | Publisher Full Text
14. Dittrich E, Raviv TR, Kasprian G, *et al.*: **A spatio-temporal latent atlas for semi-supervised learning of fetal brain segmentations and morphological age estimation.** *Med Image Anal.* 2014; **18**(1): 9–21.
PubMed Abstract | Publisher Full Text
15. Habas PA, Kim K, Corbett-Detig JM, *et al.*: **A spatiotemporal atlas of MR intensity, tissue probability and shape of the fetal brain with application to segmentation.** *NeuroImage.* 2010; **53**(2): 460–470.
PubMed Abstract | Publisher Full Text | Free Full Text
16. Makropoulos A, Counsell SJ, Rueckert D: **A review on automatic fetal and neonatal brain MRI segmentation**. *NeuroImage.* 2018; **170**: 231–248.
PubMed Abstract | Publisher Full Text
17. Serag A, Kyriakopoulou V, Rutherford M, *et al.*: **A multi-channel 4D probabilistic atlas of the developing brain: application to fetuses and neonates.** *Annals of the BMVA.* 2012; **2012**(3): 1–14.
Reference Source
18. Evans AC, Janke AL, Collins DL, *et al.*: **Brain templates and atlases.** *NeuroImage.* 2012; **62**(2): 911–922.
PubMed Abstract | Publisher Full Text
19. Dittrich E, Riklin-Raviv T, Kasprian G, *et al.*: **Learning a spatio-temporal latent atlas for fetal brain segmentation**. In: *Proceedings of the MICCAI 2011 Workshop on Image Analysis of Human Brain Development (IAHBD 2011)*; 2011.
Reference Source
20. 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 Abstract | Publisher Full Text | Free Full Text
21. Zhan J, Dinov ID, Li J, *et al.*: **Spatial-temporal atlas of human fetal brain development during the early second trimester.** *NeuroImage.* 2013; **82**: 115–126.
PubMed Abstract | Publisher Full Text | Free Full Text
22. Wu J, Sun T, Yu B, *et al.*: **Age-specific Structural Fetal Brain Atlases Construction and Cortical Development Quantification for Chinese Population.** *NeuroImage.* 2021; **241**: 118412.
PubMed Abstract | Publisher Full Text
23. Payette K, de Dumast P, Kebiri H, *et al.*: **An automatic multi-tissue human fetal brain segmentation benchmark using the fetal tissue annotation dataset.** *Sci Data.* 2021; **8**(1): 167.
PubMed Abstract | Publisher Full Text | Free Full Text
24. Doel T, Shakir DI, Pratt R, *et al.*: **GIFT-Cloud: A data sharing and collaboration platform for medical imaging research.** *Comput Methods Programs Biomed.* 2017; **139**: 181–190.
PubMed Abstract | Publisher Full Text | Free Full Text
25. Payette K, Li H, de Dumast P, *et al.*: **Fetal Brain Tissue Annotation and Segmentation Challenge Results.** arXiv preprint arXiv:220409573. 2022.
Publisher Full Text
26. Fidon L, Aertsen M, Emam D, *et al.*: **Label-set Loss Functions for Partial Supervision: Application to Fetal Brain 3D MRI Parcellation**. arXiv preprint arXiv:210703846. 2021.
Reference Source
27. Fidon L, Aertsen M, Mufti N, *et al.*: **Distributionally Robust Segmentation of Abnormal Fetal Brain 3D MRI**. arXiv preprint arXiv:210804175. 2021.
Publisher Full Text
28. Fidon L, Aertsen M, Kofler F, *et al.*: **A Dempster-Shafer approach to trustworthy AI with application to fetal brain MRI segmentation.** arXiv preprint arXiv:220402779. 2022.
Publisher Full Text
29. Ranzini M, Fidon L, Ourselin S, *et al.*: **MONAIfbs: MONAI-based fetal brain MRI deep learning segmentation**. arXiv preprint arXiv:210313314. 2021.
Publisher Full Text
30. 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 Abstract | Publisher Full Text | Free Full Text
31. Aertsen M, Verduyckt J, De Keyzer F, *et al.*: **Reliability of MR Imaging-Based Posterior Fossa and Brain Stem Measurements in Open Spinal Dysraphism in the Era of Fetal Surgery.** *AJNR Am J Neuroradiol.* 2019; **40**(1): 191–198.
PubMed Abstract | Publisher Full Text | Free Full Text
32. Garel C, Chantrel E, Brisse H, *et al.*: **Fetal cerebral cortex: normal gestational landmarks identified using prenatal MR imaging.** *AJNR Am J Neuroradiol.* 2001; **22**(1): 184–189.
PubMed Abstract | Free Full Text
33. Geerdink N, van der Vliet T, Rotteveel JJ, *et al.*: **Essential features of Chiari II malformation in MR imaging: an interobserver reliability study--part 1.** *Childs Nerv Syst.* 2012; **28**(7): 977–985.
PubMed Abstract | Publisher Full Text | Free Full Text
34. Yushkevich PA, Gao Y, Gerig G: **ITK-SNAP: an interactive tool for semi-automatic segmentation of multi-modality biomedical images**. *Annu Int Conf IEEE Eng Med Biol Soc.* IEEE; 2016; 3342–3345.
PubMed Abstract | Publisher Full Text | Free Full Text
35. Fonov V, Evans AC, Botteron K, *et al.*: **Unbiased average age-appropriate atlases for pediatric studies.** *NeuroImage.* 2011; **54**(1): 313–327.
PubMed Abstract | Publisher Full Text | Free Full Text
36. Grabner G, Janke AL, Budge MM, *et al.*: **Symmetric atlasing and model based segmentation: an application to the hippocampus in older adults.** In: *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer; 2006; 58–66.
Publisher Full Text

37. Glatter S, Prayer D, Gruber G, *et al.*: **Improved prognostication in isolated callosal agenesis: fetal magnetic resonance imaging-based scoring system.** *Ultrasound in Obstetrics & Gynecology: the Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology.* 2020.

38. Kasprian G, Langs G, Brugger PC, *et al.*: **The prenatal origin of hemispheric asymmetry: an in utero neuroimaging study.** *Cereb Cortex.* 2011; **21**(5): 1076–1083.
PubMed Abstract | Publisher Full Text

39. Gower JC: **Generalized procrustes analysis.** *Psychometrika.* 1975; **40**(1): 33–51.
Publisher Full Text

40. Gower JC: **Procrustes methods.** *Wiley Interdisciplinary Reviews: Computational Statistics.* 2010; **2**(4): 503–508.

41. Modat M, Ridgway GR, Taylor ZA, *et al.*: **Fast free-form deformation using graphics processing units.** *Comput Methods Programs Biomed.* 2010; **98**(3): 278–284.
PubMed Abstract | Publisher Full Text

42. Ashburner J, Friston K: **Non-linear registration**. Statistical parametric mapping: The analysis of functional brain images. 2007; 63–80.

43. Cachier P, Bardinet E, Dormont D, *et al.*: **Iconic feature based nonrigid registration: the PASHA algorithm.** *Computer vision and image understanding.* 2003; **89**(2-3): 272–298.
Publisher Full Text

44. Modat M, Cash DM, Daga P, *et al.*: **Global image registration using a symmetric block-matching approach.** *J Med Imaging (Bellingham).* 2014; **1**(2): 024003.
PubMed Abstract | Publisher Full Text | Free Full Text

45. Fidon L, Aertsen M, Shit S, *et al.*: **Partial supervision for the FeTA challenge 2021.** arXiv preprint arXiv:211102408. 2021.
Publisher Full Text

46. Dice LR: **Measures of the amount of ecologic association between species.** *Ecology.* 1945; **26**(3): 297–302.
Publisher Full Text

47. Fidon L, Li W, Garcia-Peraza-Herrera LC, *et al.*: **Generalised wasserstein dice score for imbalanced multi-class segmentation using holistic convolutional networks**. In: *International MICCAI Brainlesion Workshop.* Springer; 2017; 64–76.
Publisher Full Text

48. Hausdorff F: **Set theory.** *American Mathematical Society (RI).* 1991.

49. Fidon L, Viola E, Mufti N, *et al.*: **Spina Bifida Aperta Spatio-temporal Brain MRI Atlas**. *Zenodo.* 2021.
http://www.doi.org/10.5281/zenodo.5524312

50. Studholme C, Kroenke CD, Dighe M: **Motion corrected MRI differentiates male and female human brain growth trajectories from mid-gestation.** *Nat Commun.* 2020; **11**(1): 3038.
PubMed Abstract | Publisher Full Text | Free Full Text

51. Fidon L, Aertsen M, Emam D, *et al.*: **Label-set Loss Functions for Partial Supervision: Application to Fetal Brain 3D MRI Parcellation [Data set].** *Zenodo.* 2021.
http://www.doi.org/10.5281/zenodo.6878474

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 12 September 2022

https://doi.org/10.21956/openreseurope.16291.r29995

© **2022 Auzias G.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Guillaume Auzias**
Institute for Neurosciences of La Timone, UMR 7289, Aix-Marseille University, Marseille, France

The authors carefully addressed all my concerns.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Anatomical MRI data processing and analysis, Fetal MRI

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 06 April 2022

https://doi.org/10.21956/openreseurope.14998.r28870

© **2022 Auzias G.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

? **Guillaume Auzias**
[1] Institute for Neurosciences of La Timone, UMR 7289, Aix-Marseille University, Marseille, France
[2] Institute for Neurosciences of La Timone, UMR 7289, Aix-Marseille University, Marseille, France

This work introduces a spatio-temporal atlas specifically designed for fetuses suffering from spina bifida aperta, enabling improved automatic segmentation compared to using an atlas of normal fetuses.

The description of the entire study is quite complete, with extensive information provided such as the protocol for anatomical landmarks annotation. The proposed atlas is available and easily accessible following the link provided in the article. The efforts for sharing the data and tools are much appreciated.

While the method is well described and solid, I have few important concerns regarding limitations related to the data available. Also, the motivations for some aspects of the approach are not sufficiently detailed. I believe that addressing my concerns and comments below would clearly benefit the final version of this paper, and would thus contribute to the dissemination of the atlas and tools released.

Major concern 1

The number of subjects involved for the definition of each age-specific atlas can be as small as 3, which is likely to be too small to account for the inter-individual variability. Given the high variability of the brain anatomy in SBA as acknowledged by the authors, 37 subjects might be insufficient to cover the whole spectrum of inter-individual variability and/or the whole spectrum of variations induced by the pathology. More specifically, the atlases at 33 and 34 weeks show a poor delineation and segmentation of the cortical gray matter, which is likely due to the lack of data available to compute the average image. An underestimation of the true inter-individual variance is also possible even when the atlas is sharp. While this is less of a problem for the delineation of the cortical gray matter that is very smooth at early developmental stage, it could affect the delineation of the other structures of great interest such as the ventricles and the cerebellum. This also likely affects the automatic segmentation and might explain the modest performances in particular for the cortical gray matter with a Dice inferior to 50. This is an important limitation for future applications of this atlas that should be explicitly acknowledged and stated in section 7.

Major concern 2

In the proposed approach, the age-specific atlases are built by grouping the fetuses based on their gestational age. This approach suffers from limitations such as the dependency to the number of available cases per age as pointed in my previous concern. Why didn't you consider using time-varying kernel as proposed in (Serag *et al.*, ref 16 of the submission) to overcome the variations in the distribution of subjects at different ages? I acknowledge this would add some complexity to the weighted generalized Procrustes method described in section 3.2.2, but a sliding window approach as proposed in Serag *et al.* could work, no? More generally, please refer explicitly to Serag *et al.* in Section 3.2 and clarify how your approach relates to that one, that is probably the most closely related publication.

Major concern 3

The sex of fetuses is not considered in the present study. The recent publication (Studholme *et al.*, 2020, see complete ref below) reported different brain growth trajectories between male and female fetuses. The potential implications for the process of atlas building are not straightforward. A study of the influence of sex distribution on such an atlas would be relevant, but I acknowledge it might fall out of the scope of current submission. If the information of fetus sex was recorded,

please report the distributions, and if it was not recorded, please mention this absence of information as a limitation the section 7.

Colin Studholme, Christopher D. Kroenke, and Manjiri Dighe, “Motion Corrected MRI Differentiates Male and Female Human Brain Growth Trajectories from Mid-Gestation,” Nature Communications 11, no. 1 (June 16, 2020): 3038, https://doi.org/10.1038/s41467-020-16763-y.

Major concern 4

The motivations for using anatomical landmarks to constraint the registration are not clear.
Is the fully automated intensity-driven registration failing? Did you try to adapt the registration algorithm parameters such as e.g. the multi-resolution pyramid? Please add a subsection to state how and why the classical intensity-driven approaches fail in this context.

Major concern 5

Regarding the experiment on the automatic segmentation based on the SBA-specific atlas, an important information is lacking: are the landmarks used for registration when the SBA-specific atlas is used? If the answer is yes, then the comparison with the atlas of normal fetuses is unfair, and an additional experiment without the use of the landmarks with the SBA-specific atlas would be welcome.

Minor concern 1

Make more explicit that steps 3.1.1, 3.1.2 and 3.1.3 of the proposed pipeline are achieved by NiftyMIC on Fig2, e.g. by using a specific color.

Minor concern 2

Please add a small subsection in 3.1.5 providing supplementary information regarding the impact of the surgery on the anatomy of fetuses and comment on how these changes are expected to influence the resulting atlas.

Minor concern 3

In section 3.2.1, the sigma of the Gaussian kernel is set to a value of 3 without any justification.
Please precise how this parameter value was set and motivate the (expected) limited influence of that parameter on the resulting atlas.

Minor concern 4

An artefact is visible on the 27 week atlas in the orbito-frontal region.
This artefact does only affect the atlas at that particular age.
Please investigate and comment.

Minor concern 5

The reference to Table 1 is missing in the text of section 5.1.

Other minor concerns / Typos

In the Plain language summary:
“Visualization” is appropriate? -> Understanding?
“Expected” -> normal?

In the introduction:
“SBA fetuses have also smaller hippocampus7, abnormal cortical thickness and gyrification8,9 , and smaller deep grey matter volume and total brain volume10,11”. References 7, 9, 10, 11 actually involve pediatric or adult populations, while this sentence suggest that all these results were obtained in fetuses, which is thus incorrect. Please be more specific in separating the studies on fetuses or on post-natal data.

“Atlases can also be used to measure variability in the brain anatomy of an individual as compared to the whole population” to the whole population -> to the model supposed to be representative of the whole population.

Those anatomical landmarks are used for two things in our pipeline -> please rephrase.

References
1. Studholme C, Kroenke C, Dighe M: Motion corrected MRI differentiates male and female human brain growth trajectories from mid-gestation. *Nature Communications*. 2020; **11** (1). Publisher Full Text

Is the rationale for developing the new method (or application) clearly explained?
Yes

Is the description of the method technically sound?
Partly

Are sufficient details provided to allow replication of the method development and its use by others?
Partly

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Anatomical MRI data processing and analysis, Fetal MRI

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 24 Aug 2022

Lucas Fidon

We thank the reviewer Guillaume Auzias for his insights and comments. They have allowed us to improve the manuscript and our atlas. We respond below to the concerns of the reviewer one by one.

Major concern 1 While limited, the current number of subjects included in the atlas construction already allowed us to reach useable performance as illustrated in our evaluation section. Increasing this number would obviously be of interest but was found to be impractical within the scope of this project. From a methodological perspective, we have mitigated the low number of subjects for some gestational ages using Gaussian kernel regression (please see equation (1) in the revision) and right-left symmetry (please see 3.1.5 Data augmentation in the revision). We give more details about this in our response to major concern 2 below. To better analyze the segmentation performance per gestational age, we have added the fetal brain 3D MRIs from the FeTA dataset release 2 to the segmentation evaluation cohort. This brings the total of MRIs of neurotypical (resp. spinal bifida) cases used in our evaluation to 32 3D MRIs (resp. 36 3D MRIs). Results in Table 2 in the revision have been updated and confirm that the proposed fetal brain atlas for spina bifida outperforms a state-of-the-art neurotypical fetal brain atlas in terms of mean Dice scores for all tissue types for spinal bifida cases. In addition, we have added two new figures (see Fig. 9 and Fig. 10 in the revision) with boxplots of the Dice scores and Hausdorff distance for each method, each tissue type, and each gestational age (in weeks) separately. Those figures are only for the spina bifida cases of the evaluation cohort (n=36). Regarding the cortical gray matter of spina bifida cases specifically, the boxplots show that the Dice scores are lower and the Hausdorff higher at earlier developmental gestational ages. This suggests that contrary to the hypothesis of the reviewer, the smoothness of the cortical gray matter is not the anatomical feature dominating the difficulty of the automatic segmentation of the cortical gray matter. Besides, the proposed fetal brain atlas for spina bifida outperforms the neurotypical atlas for the cortical gray matter for two segmentation metrics and all gestational age, except for 27 weeks. It is worth noting that the gestational ages range for evaluation is 20 weeks to 30 weeks for spina bifida. This is because there are no MRIs of spina bifida cases older than 30 weeks at the time of acquisition in the FeTA dataset. In particular, the atlases at 33 and 34 weeks have no influence on the Dice scores reported in the paper. We have clarified this in the dataset description. This is also a limitation of our evaluation and we have added it in the limitations section. We agree with the reviewer that the cortical segmentation at 33 and 34 weeks was rather poor, especially at the insula in both weeks and in the occipital and inferotemporal cortical lining at 34 weeks. Co-author Michael Aertsen has now corrected these segmentations and we have updated the atlas online. We do acknowledge that our open-source segmentations for the proposed spina bifida atlas might remain suboptimal for some applications. We would like to emphasize that we have made the atlas and the segmentation available under the license

CC0 1.0 Public domain dedication to allow other researchers to improve and redistribute the segmentations if they wish.

Major concern 2 We apologize for the confusion and clarify that, similar to Serag et al., we already use time-varying kernels in our approach. The time-varying kernel is defined in equation (1) of paragraph 3.2.1 in the revision. The time-varying weights defined in (1) are used at two stages of the computation of the atlas:
1. for the initialization of the age-specific atlases in the generalized Procrustes method (see equations (4), (5), and (6))
2. in every update of the age-specific atlases (see equation (2)). We think paragraph 3.1.5 in the initial manuscript about the age-specific groups might have been confusing. We have clarified the paragraph: "Each group is assigned with a gestational age ranging from 21 weeks to 34 weeks. Volumes are included in a group only if the gestational age at the time of the acquisition is within 9 days of the gestational age of the group." By adding: "This implies that there are overlaps between groups. For example, the 24 weeks group contains the fetal brain MRIs acquired between 22 weeks + 4 days and 25 weeks + 3 days of gestation. In addition, the contribution of each volume within an age-specific group is weighted using a time-varying Gaussian kernel, as defined in the next section in (1). The value of 9 days, used above, is chosen to correspond to 3 x sigma where sigma is defined in the time-varying Gaussian kernel regression (1)." To improve clarity, in section titles and in fig. 2, we have also replaced "weighted" by "time-weighted". We have also added a citation to Serag et al before equation (1) in the revision.

Major concern 3 A figure showing the distribution per gestational age of the genders of the fetuses whose 3D MRIs were used to compute the atlas has been added in Fig. 1 (right part) of the revision. We have added in the limitations: "The proposed atlas mixes male and female fetuses. However, recent work reported different brain growth trajectories between male and female neurotypical fetuses (Studholme et al., 2020). Sex-specific atlases may be of interest to better represent the populations of male and female fetuses with spina bifida aperta." We thank the reviewer for the suggestion. Despite only females being present between 28 weeks and 30 weeks of gestation, globally we have found no statistical difference between the distributions of gestational ages for males and females using a Mann-Whitney U test with a confidence level of 95% (p-value=0.35).

Major concern 4 We have several motivations that support the use of the proposed anatomical landmarks for computing the atlases. Firstly, without an informative initialization, we found that intensity-driven registration cannot cope with the large variability observed in fetal imaging for spinal bifida. Then, the Procrustes method used for the initialization of the atlases could not be performed without the anatomical landmarks since it uses only the landmarks and not the intensity (please see equations (3) to (6) in the revision). Secondly, in the paragraph parameter tuning of section 3.2.3 in the revision, we give details of the grid search of the hyper-parameters that we have conducted for the non-linear registration. We conclude this paragraph with "Our parameter tuning protocol suggests that all the terms of the objective function are important to obtain optimal image registration results." This includes the term related to the anatomical landmarks. We have clarified this by adding: "In particular, this supports the usefulness of the landmarks for the registration since a non-minimal value of alpha_LMKS was optimal."

Major concern 5 We considered that our landmarks would not typically be available at test time for the purposes of atlas-based segmentation. The landmarks are thus not used during registration for the evaluation of automatic segmentation. We have clarified this in section 5.2 of the revision.

Minor concern 1 We have updated Fig 2 in the revision as requested.

Minor concern 2 We have added a paragraph about this in the discussion: "With fetal surgery, the open neural tube defect is closed and thus the continuous leakage of cerebrospinal fluid (CSF) is stopped. This leads to an increase in CSF within the skull, surrounding the cerebrum and cerebellum, leading to a better distinction of the grey matter from the inner lining of the skull. This may explain the higher segmentation accuracy for the grey matter after 27 weeks as from this timepoint the evaluation cohort includes only postoperative images. In addition to the increase in CSF within the skull, the closure of the defect leads to a reversal of the hindbrain herniation, as demonstrated in a previous paper by our group this happens already within 7 days after surgery in the majority of cases (Aertsen et al, 2019 AJNR). The reversal of the hindbrain information in combination with the increase in fluid surrounding the cerebellum and brainstem in the posterior fossa, improves the distinction of the cerebellum and brainstem from the skull base. Again we notice an improved performance after 27 weeks confirming the impact of fluid restoration in the skull on our automatic segmentation algorithm."

Minor concern 3 We have chosen the value sigma = 3 days so that an interval [-sigma, sigma] covers approximately one week which is the time unit for the atlases. We have not tested other values of sigma and we have not evaluated the influence of sigma on the resulting atlas. This is now clarified in the revision, please see 3.2.1.

Minor concern 4 We think that this artefact is due to variability in the topology of the extra-axial cerebrospinal fluid (CSF) at week 26 weeks for operated fetuses. At this time point, the extra-axial CSF can be either circumferential or limited as illustrated in Fig. 14 that we have added in the revision. In this case, the diffeomorphic hypothesis of our non-linear registration step is violated and can lead to such artefacts. It is worth noting that this issue affects the atlas at 27 weeks due to the use of a time-weighted regression kernel. We have investigated the segmentation performance per tissue type and per gestational age in Fig 9 and Fig 10 of the revision. Week 27 is the only week for which the proposed atlas underperforms the baseline for the extra-axial CSF and the deep grey matter. The artefact in the orbito-frontal region may account for the suboptimal automatic segmentation at 27 weeks. In addition, it is worth noting the large variability of the segmentation results for week 26 in Fig. 9 and Fig.10 of the revision for the two atlases. This can be attributed to the variability in the topology of the extra-axial CSF illustrated in Fig. 14 of the revision. At week 26, the spina bifida atlas performs best on 3D MRIs of fetuses with limited extra-axial CSF (Fig. 14 of the revision left) while the neurotypical atlas performs best for 3D MRIs of fetuses with circumferential extra-axial CSF (Fig. 14 of the revision right). We have added the comments above to the discussion and the limitations.

Minor concern 5 We have added the missing reference in section 5.1 of the revision.

Other minor concerns / Typos Typos were corrected, thank you. We have clarified the use of some references in the introduction: "In postnatal life, children and adults with spina bifida aperta are known to have smaller hippocampus7, abnormal cortical thickness and gyrification8,9 , and smaller deep grey matter volume and total brain volume10,11. In a small pilot study, it has been observed that fetal brain volume and shape is different after spina bifida repair compared to controls (Mufti et al, neuroradiology, 2021)"

Competing Interests: No competing interests were disclosed.

Reviewer Report 26 November 2021

https://doi.org/10.21956/openreseurope.14998.r27829

© 2021 Jakab A. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Andras Jakab
[1] Center for MR-Research, University Children's Hospital Zürich, Zürich, Switzerland
[2] Center for MR-Research, University Children's Hospital Zürich, Zürich, Switzerland

This article describes the first publicly available spatio-temporal fetal cerebral MRI atlas for spina bifida. The atlas comprises anatomical templates and multi-tissue segmentation maps for different gestational ages and operated vs. non-operated fetuses. The manuscript adequately describes the methods that were used to create the dataset. One novelty is that the work used anatomical landmarks and a weighted generalized Procrustes method for the temporal regularization of data. I believe that the atlas will be a valuable resource for the quantitative, automated and large-scale analysis of SBA MRI data.

The data is high quality and the methods are reproducible. As the authors point out, the work has a few weaknesses. For example, the atlas is based on low case numbers for the more advanced pregnancies. To cover this age range, the authors may consider including postnatal images since many of these infants are born prematurely. SBA fetuses have a large variability in terms of their ventricular width. In the clinical practice, cases with much larger ventricles as the mean template in the Atlas are common. The analysis of these cases might be problematic. Perhaps in the future, the authors could also consider generating anatomical templates with a varying level of hydrocephalus.

I would suggest that the authors consider citing the final FeTA article instead of the pre-print: https://doi.org/10.1038/s41597-021-00946-3

References

1. Payette K, de Dumast P, Kebiri H, Ezhov I, et al.: An automatic multi-tissue human fetal brain segmentation benchmark using the Fetal Tissue Annotation Dataset. *Scientific Data*. 2021; **8** (1).

Publisher Full Text

Is the rationale for developing the new method (or application) clearly explained?
Yes

Is the description of the method technically sound?
Yes

Are sufficient details provided to allow replication of the method development and its use by others?
Yes

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Fetal and infant neuroimaging research. Group leader / PI.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 24 Aug 2022

Lucas Fidon

We thank the reviewer András Jakab for assessing our work and for his feedback. We have updated the reference to the FeTA dataset as suggested.

Competing Interests: No competing interests were disclosed.