



OPEN

A Fetal Brain magnetic resonance Acquisition Numerical phantom (FaBiAN)

Hélène Lajous^{1,2}✉, Christopher W. Roy^{1,7}, Tom Hilbert^{1,3,4,7}, Priscille de Dumast^{1,2}, Sébastien Tourbier¹, Yasser Alemán-Gómez¹, Jérôme Yerly^{1,2}, Thomas Yu⁴, Hamza Kebiri^{1,2}, Kelly Payette^{5,6}, Jean-Baptiste Ledoux^{1,2}, Reto Meuli¹, Patric Hagmann¹, Andras Jakab^{5,6}, Vincent Dunet¹, Mériam Koob¹, Tobias Kober^{1,3,4,7}, Matthias Stuber^{1,2,7} & Meritxell Bach Cuadra^{1,2}

Accurate characterization of in utero human brain maturation is critical as it involves complex and interconnected structural and functional processes that may influence health later in life. Magnetic resonance imaging is a powerful tool to investigate equivocal neurological patterns during fetal development. However, the number of acquisitions of satisfactory quality available in this cohort of sensitive subjects remains scarce, thus hindering the validation of advanced image processing techniques. Numerical phantoms can mitigate these limitations by providing a controlled environment with a known ground truth. In this work, we present FaBiAN, an open-source Fetal Brain magnetic resonance Acquisition Numerical phantom that simulates clinical T2-weighted fast spin echo sequences of the fetal brain. This unique tool is based on a general, flexible and realistic setup that includes stochastic fetal movements, thus providing images of the fetal brain throughout maturation comparable to clinical acquisitions. We demonstrate its value to evaluate the robustness and optimize the accuracy of an algorithm for super-resolution fetal brain magnetic resonance imaging from simulated motion-corrupted 2D low-resolution series compared to a synthetic high-resolution reference volume. We also show that the images generated can complement clinical datasets to support data-intensive deep learning methods for fetal brain tissue segmentation.

Today, there is a growing awareness of the importance of early brain development on health later in life^{1–9} as brain maturation involves complex and interconnected structural and functional processes that can be altered by various genetic and environmental factors. Magnetic resonance imaging (MRI) may be required during pregnancy to investigate equivocal situations as a support for diagnosis and prognosis, but also for postnatal management planning¹⁰. In clinical routine, T2-weighted (T2w) fast spin echo (FSE) sequences are used to scan multiple 2D thick slices that provide information on the whole brain volume with a good signal-to-noise ratio (SNR) while minimizing the effects of random fetal motion during acquisition¹¹. However, stochastic movements of the fetus in the womb cause various artefacts in the images, including drops in signal intensity. Post-processing approaches built on motion estimation and correction can compensate for such artefacts. Especially, super-resolution (SR) reconstruction techniques take advantage of the redundancy between low-resolution (LR) series acquired in orthogonal orientations to reconstruct an isotropic high-resolution (HR) volume of the fetal brain with reduced intensity artefacts and motion sensitivity^{12–17}. Navigating through the resulting SR volume provides valuable information on the developing brain anatomy, including consistent biometric measurements^{18–20}. Besides, fetal brain tissue segmentation is critical for further investigation of brain development, especially for volumetric evaluation^{21–25}.

Manual segmentation is a cumbersome and time-consuming task. Therefore, supervised deep learning approaches that rely on annotated data have emerged as accurate techniques for automated delineation of the

¹Department of Radiology, Lausanne University Hospital (CHUV) and University of Lausanne (UNIL), Lausanne, Switzerland. ²CIBM Center for Biomedical Imaging, Lausanne, Switzerland. ³Advanced Clinical Imaging Technology (ACIT), Siemens Healthcare, Lausanne, Switzerland. ⁴Signal Processing Laboratory 5 (LTS5), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. ⁵Center for MR Research, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland. ⁶Neuroscience Center Zurich, University of Zurich, Zurich, Switzerland. ⁷These authors contributed equally: Christopher W. Roy, Tom Hilbert, Tobias Kober and Matthias Stuber. ✉email: helene.lajous@unil.ch

fetal brain^{22–24,26}. The development and validation of such advanced image processing and analysis methods require access to large-scale data to account for the subject variability, but the number of good quality, exploitable MR acquisitions available in this sensitive cohort remains relatively scarce.

Numerical simulations can mitigate these limitations by providing a controlled environment with a known ground truth for accurate, robust and reproducible research^{27,28}. MR developments often rely on computer simulations that enable pulse sequence design, accurate prototyping and evaluation of new advanced acquisition schemes as well as validation of reconstruction techniques in a controlled setting^{27–31}. Such platforms are also valuable educational tools for physicists and technologists^{29,30}. In this sense, MR simulations can efficiently complement or even replace the design and use of sophisticated experimental phantoms, as well as experiments on animal models or even on human volunteers³⁰. MR simulators can be designed to address multiple challenges, such as system imperfections, multichannel transmission, correction/suppression of image artefacts, and optimization of specific absorption rate (SAR)^{31,32}.

Motion is a major hurdle in various MRI applications, from cardiovascular MR to functional MRI analysis and fetal imaging, as it is responsible for artefacts in the images and can lead to erroneous data analysis and interpretation^{27,28,30,32}. Whereas periodic movements can be directly related to physiological processes such as breathing or a heartbeat, and may therefore be compensated during post-processing, stochastic fetal motion impedes the repeatability of measurements²⁸ and thus hinders retrospective motion correction. The difficulty of estimating such unpredictable movements results in the lack of any ground truth, yet necessary for the validation of new methods³². Numerical phantoms are an interesting alternative that offers a fully scalable and flexible environment where any image-acquisition, -reconstruction, or -processing technique can be evaluated, optimized, and validated from a collection of synthetic, yet realistic data that simulate multiple controlled conditions. Furthermore, the results obtained by these various strategies can be quantitatively compared to each other through simulated reference data based on full-reference image quality assessment metrics such as the mean squared error (MSE), the peak signal-to-noise ratio (PSNR), or even the more perceptual structural similarity index (SSIM)^{12,33,34}.

Paradoxically, the diversity of MRI simulators, each with its own advantages and limitations, jeopardizes the comparison from one setup to another²⁷. In contrast to a simplified analytical description of the MR signal arising from proton isochromats, advanced developments in the field of MRI require more realistic numerical simulations³¹. Two main approaches have been investigated: on one hand, (i) the analytical numerical formalism is based on a mathematical description of both the anatomy and the MR experiment as in the Shepp–Logan head phantom³⁵, and on the other hand, (ii) voxel-based phantoms are usually derived from segmented clinical acquisitions relevant to the targeted application²⁷. The correspondence between the image and the corresponding k-space is governed by the continuous Fourier transform in analytical numerical models which enables an accurate representation of k-space, whereas it is approximated by its discrete version in their voxel-based equivalents. Analytical numerical phantoms are powerful tools to study k-space truncation artefacts while voxel-based simulations can be used to also model physiological processes and motion that may alter the acquisitions^{27,28}. However, spatial and temporal resolutions depend on the original images from which voxel-based phantoms are derived, which may limit their use in reproducibility studies.

Hybrid phantoms have been developed to leverage both approaches and overcome their respective limitations, resulting in versatile and realistic models^{27,28}. Depending on the targeted application, advanced phantoms can be built on these models to include more features as in the case of the fetal extended Cardiac-Torso (XCAT) cardiovascular magnetic resonance imaging (Fetal XCMR) phantom that combines two independent XCAT models of both the anatomy and physiology of a mother and her baby with a simulation framework for 2D cardiovascular magnetic resonance (CMR) acquisitions²⁸. Of note, realistic MRI simulations are hampered by the high computational burden associated with the 3D representation of the simulated object and the introduction of motion during acquisition^{30,31}. Although the parallelization of calculations on computer clusters allows to speed up the simulations, such advanced infrastructures are not always available.

Various advanced MRI simulation platforms^{31,36,37} are based on the numerical resolution of Bloch equations which may be demanding due to the need to compute a numerical solution for each resonant frequency within a single voxel. The extended phase graph (EPG) concept^{38–40} is a surrogate for Bloch equations to describe the magnetization response to various MR pulse sequences, including complex acquisition schemes that involve multiple radiofrequency (RF) pulses and gradients. Modeling the evolution of spin magnetization depending on tissue properties allows to gain insight into the obtained MR signal and to evaluate its behavior. Originally suggested to assess signal intensities in multi-echo experiments with variable flip angles⁴¹, the EPG algorithm has aroused growing interest in recent years, especially for precise characterization of echoes^{42–44} and diffusion effects⁴⁵, evaluation of physics-constrained reconstruction methods⁴⁶, sequence pulse design^{38,47,48} and quantitative MRI techniques^{49–51}. The EPG formalism relies on the Fourier representation of the evolution of spin magnetization within a voxel after application of various RF pulses and gradients. It assumes that a single set of relaxation parameters can characterize a given tissue⁴⁰. The Fourier series are used to account for the multiple resonant frequencies that may arise from local magnetic field inhomogeneities within a voxel. Thus, the EPG algorithm is able to provide fast and accurate simulations.

To our knowledge, there is no simulation framework for fetal brain MRI. In this work, we present FaBiAN, an open-source Fetal Brain magnetic resonance Acquisition Numerical phantom that simulates T2w images of the in utero developing brain built on segmented anatomical images from a previously published normative spatiotemporal HR MRI atlas of the fetal brain¹⁰. FaBiAN relies on the EPG formalism^{38–40} of the signal formation to simulate FSE acquisitions of the fetal brain, in this case Half-Fourier Acquisition Single-shot Turbo spin Echo (HASTE, Siemens Healthineers) and Single-Shot Fast Spin Echo (SS-FSE, GE Healthcare) sequences, based on a flexible and realistic setup that accounts for intensity non-uniformity fields and stochastic fetal motion. We investigate the capabilities of the developed framework and provide a proof of concept of its practical value

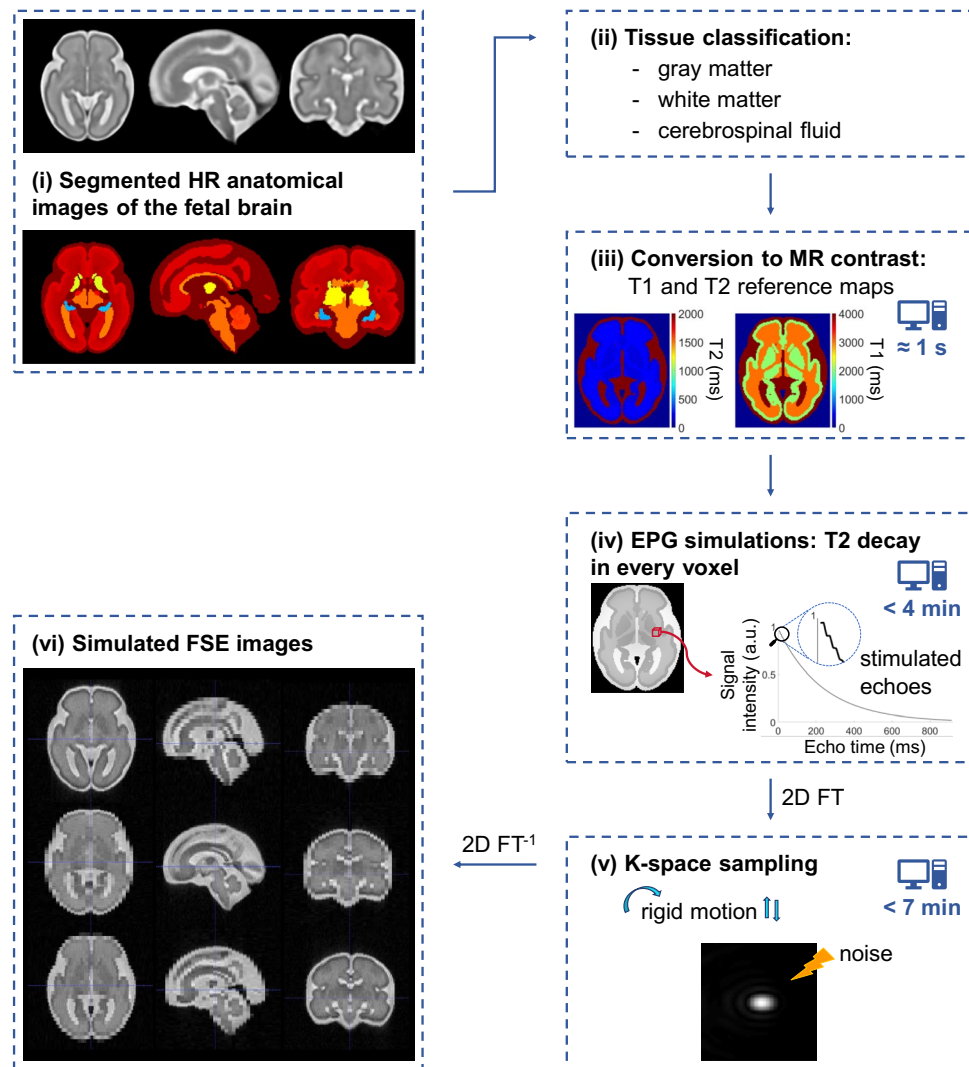


Figure 1. Workflow for simulating images of the fetal brain acquired by a fast spin echo (FSE) sequence (i) from segmented HR anatomical MR images¹⁰, illustrated for a fetus of 30 weeks of GA. (ii) Brain tissues are classified into gray matter, white matter and cerebrospinal fluid. (iii) Anatomical structures are converted to the corresponding MR contrast to obtain reference T1 and T2 maps of the fetal brain at either 1.5 or 3 T. (iv) The EPG algorithm allows to accurately simulate the T2 decay over time in every brain voxel by accounting for the effects of the stimulated echoes, as highlighted by the enlargement of the beginning of the curve. This spatiotemporal information is subsequently used (v) to sample the Fourier domain of the simulated images of the moving fetus. After the addition of noise to match the SNR of real clinical acquisitions, (vi) FSE images of the fetal brain are eventually recovered by 2D inverse Fourier transform ($2D FT^{-1}$).

in two key application examples. First, we generate LR images with multiple levels of motion to fine-tune a SR reconstruction algorithm^{16,52}. Then, we explore the potential of using multiple fetal brain images simulated at different gestational ages (GA) with several motion amplitudes and a variety of acquisition parameters for data augmentation in fetal brain tissue segmentation.

Methods

Numerical implementation of FSE sequences. Figure 1 provides an overview of the workflow implemented in MATLAB (MathWorks, R2019a) to simulate T2w fetal brain images acquired using clinical FSE sequences. (i) High-resolution anatomical images from a normative spatiotemporal MRI atlas¹⁰ are used as a model of normal fetal brain. (ii) Segmented brain tissues are organized into gray matter, white matter and cerebrospinal fluid as shown in Table 1, and (iii) are assigned relaxometry properties from the literature at 1.5 T^{53–57} or 3 T^{58–62} accordingly. (iv) The T2 decay over time is computed in every voxel of the HR anatomical images from the sequence parameters using the EPG formalism^{39,40}. (v) The Fourier domain, or k-space, of the simulated images is sampled from the T2 decay matrix to reflect the process of FSE acquisition in the presence of random rigid motion. (vi) The final simulated images are recovered by a 2D inverse Fourier transform.

Gray matter	White matter	Cerebrospinal fluid
Amygdala	Cerebellum	Cerebrospinal fluid
Caudate	Corpus callosum	Lateral ventricle
Cortical plate	Fornix	
Hippocampus	Hippocampal commissure	
Putamen	Intermediate zone	
Subthalamic nuclei	Internal capsule	
Thalamus	Midbrain	
	Miscellaneous	
	Subplate	
	Ventricular zone	

Table 1. Classification of segmented brain tissues¹⁰ as gray matter, white matter and cerebrospinal fluid.

The clinical implementation of FSE sequences can vary from one MR vendor to another, but also according to the technical characteristics and hardware limitations of the MR scanner. For instance, the gradients are characterized by their slew rate that directly impacts the echo time. Especially in the context of fetal imaging, different SAR limits may be enforced, resulting in various flip angles. Furthermore, a specific clinical indication may require adjusting standard acquisition parameters, for example by increasing the effective echo time to enhance the contrast between the subplate and surrounding structures, and investigate pathologies of the cerebrospinal fluid. Therefore, we have developed this open-source numerical phantom with the idea of keeping the framework as general as possible to enable users a large flexibility in the type of images simulated. As such, multiple acquisition parameters can be set up with respect to the MR contrast (effective echo time, excitation/refocusing pulse flip angles, echo spacing, echo train length), the geometry (number of 2D slices, slice orientation, slice thickness, slice gap, slice position, phase oversampling), the resolution (field-of-view, matrix size), the resort to any acceleration technique (acceleration factor, number of reference lines) or to scanner interpolation, as well as other settings related to the age of the fetus, the RF transmit field inhomogeneities, the amplitude of random fetal motion in the three main directions, and the SNR. Table 2 provides the ranges of values used in this study to simulate T2w images of the fetal brain as they are acquired in clinical routine at two sites: Lausanne University Hospital (CHUV, HASTE sequence) and University Children's Hospital Zurich (Kispi, SS-FSE sequence). Additionally, Supplementary Table S1 reports the ranges of values that are relevant to an application in fetal brain MRI, thus consistent with common clinical protocols, and that were tested during the development of this first prototype. FaBiAN allows to transpose the main differences between the HASTE and SS-FSE protocols related to timing (effective echo time and echo spacing), geometry (slice thickness, slice gap, and phase oversampling), and spatial resolution (field-of-view and acquisition/reconstruction matrix). Irrespective of the MR vendor, the slice profile is modeled by a 2D Gaussian function with the full width at half maximum equal to the slice thickness in the slice-selection direction¹⁶. The sampling trajectory is assumed to be the same in both protocols.

The entire simulation pipeline is described in detail in the following.

Fetal brain model and MR properties. Our numerical phantom is based on segmented 0.8-mm-isotropic anatomical images (Fig. 1-(i)) from the normative spatiotemporal MRI atlas of the developing brain built by Gholipour and colleagues from normal fetuses scanned between 19 and 39 weeks of gestation¹⁰. Due to the lack for ground truth relaxometry measurements in the fetal brain, all thirty-four segmented tissues are merged into three classes according to medical experts: gray matter, white matter and cerebrospinal fluid (Fig. 1-(ii) and Table 1). Corresponding T1 and T2 relaxation times at 1.5 T^{53–57} are assigned to these anatomical structures to obtain reference T1 and T2 maps, respectively (Fig. 1-(iii)). The value of these relaxometry properties at 3 T is estimated from the literature^{58–62}, assuming that doubling the magnetic field strength from 1.5 to 3 T increases the T1 relaxation time by approximately 25% in gray matter and 10% in both white matter and cerebrospinal fluid, while the T2 properties remain unchanged.

Intensity non-uniformity (INU) fields. Non-linear slowly-varying INU fields due to transmit field inhomogeneities (B_1^+) are based on BrainWeb estimations from real scans to simulate T2w images^{63–67}. The available 20% INU version is resized to fit the dimensions of the atlas images and normalized by 1.2 to provide multiplicative fields in the range of 0.8 to 1.2 over the brain area. It is subsampled to a 0.1-mm resolution with linear interpolation in the slice thickness orientation in order to account for B1 bias field variations across the slice profile.

EPG formalism. The evolution of thousands of spin systems (i.e. isochromats) over time is commonly studied using Bloch equations to characterize echoes. However, this method is computationally expensive, because the equations must be solved for each isochromat, and not fully accurate to determine the echo intensity. Therefore, and since FSE sequences involve multiple successive RF pulses, the EPG concept seems particularly relevant to our application. Indeed, it combines the configuration state and partition state approaches to efficiently account for the evolution of a significant ensemble of isochromats, and thus accurately characterize echoes (type, inten-

Simulations	HASTE	SS-FSE
GA (weeks)	21–33	21–35
Magnetic field strength	1.5 T	1.5 T or 3 T
Field inhomogeneities	20% INU level provided by BrainWeb ^{63,64} for T2w images	
Acquisition parameters		
Contrast		
Effective echo time (ms)	90	116.256–123.60
Echo spacing (ms)	4.08	10
Echo train length	224	224
Excitation flip angle (°)	90	90
Refocusing pulse flip angle (°)	180	180
Geometry		
Slice orientation	Sagittal, coronal or transverse	Sagittal, coronal or transverse
Slice thickness (mm)	3	3–4
Slice gap (mm)	0.3	0
Number of slices	45–46	37–51
Phase oversampling (%)	80	0
Shift of the field-of-view (mm)	± 1.6	± 1.6
Resolution		
Field-of-view (mm ²)	360 × 360	240 × 240–300 × 300
Base resolution (voxels)	320–327	256
Phase resolution (%)	70	100
Reconstruction matrix	320 × 404–327 × 414	512 × 512
Zero-interpolation filling	–	Yes
Acceleration technique		
Reference lines	42	–
Acceleration factor	2	–
Amplitude of 3D rigid motion		
Little motion		
Translation (mm) in x	± 1	± 1
Translation (mm) in y	± 1	± 1
Translation (mm) in z	± 1	± 1
3D rotation (°)	± 2	± 2
Moderate motion		
Translation (mm) in x	± 3	± 3
Translation (mm) in y	± 3	± 3
Translation (mm) in z	± 3	± 3
3D rotation (°)	± 5	± 5
Strong motion		
Translation (mm) in x	± 4	± 4
Translation (mm) in y	± 4	± 4
Translation (mm) in z	± 4	± 4
3D rotation (°)	± 8	± 8
Noise		
Mean	0	0
Standard deviation	0.15	0.01

Table 2. The flexibility of FaBiAN is illustrated by the number of sequence parameters and settings available to the user. The ranges of values used to simulate fetal brain MR images are presented, in agreement with the clinical protocols respectively in place at CHUV (HASTE sequence) and Kispi (SS-FSE sequence). The differences in the implementation of both sequences mainly rely on the simulation of the GRAPPA acceleration technique for the HASTE, which affects the way the k-space of the simulated images is sampled, and the simulation of the scanner in-line interpolation for the SS-FSE, which requires low-pass filtering before zero-interpolation filling of k-space.

sity, timing)^{39,68} (see Supplementary Method S1, Supplementary Fig. S1 for further details on the EPG formalism).

The EPG algorithm³⁹ simulates the T2 decay in every voxel of the anatomical images over each echo train (Fig. 1-(iv)) based on the FSE sequence pulse design. Along with reference T1 and T2 maps of the fetal brain, the following sequence parameters are fed into the EPG estimation algorithm: the effective echo time, the excitation and refocusing pulse flip angles that modulate the realistic INU fields described above, resulting in spatially varying flip angles, and the echo train length, namely the number of repetitions of RF refocusing flip angles. The resulting 4D matrix that combines information about both the anatomy and the magnetic relaxation properties of the fetal brain is hereafter referred to as the T2 decay matrix.

K-space sampling and image formation. The T2 decay matrix is Fourier-transformed and subsequently used for k-space sampling of the simulated images. For a given echo time (TE), at most one line from the associated Fourier domain of the T2 decay matrix is used, with the central line corresponding to the effective TE. If an acceleration technique such as GRAPPA interpolation is required to decrease the scanning time as in the HASTE sequence acquired in clinical routine at CHUV, multiple reference lines are consecutively sampled around the center of k-space. Beyond, one line out of two is actually needed to simulate an acceleration factor of two. As a first approximation, these sampled lines are copied to replace the missing lines. According to partial Fourier imaging techniques, the properties of Hermitian symmetry in the frequency domain are used to fill the entire k-space. While intra-slice motion can be neglected in FSE sequences, inter-slice random 3D translation and rotation of the fetal brain are implemented during k-space sampling (Fig. 1-(v)). If needed, zero-interpolation filling (ZIP) is performed by filling the edges of the simulated k-space with zeros in order to reach the desired reconstruction matrix size, as for the high in-plane resolution SS-FSE images acquired at Kispi. Data in k-space are previously processed using a Fermi low-pass filter with a radius of 0.85 and a width of 1/23 to avoid Gibbs ringing artefacts^{69,70}. Complex Gaussian noise (mean, 0; standard deviation, 0.15 for the HASTE, 0.01 for the SS-FSE implementation respectively) is added to simulate thermal noise generated during the acquisition process and qualitatively match the SNR of clinical data. The simulated images are eventually recovered by 2D inverse Fourier transform (Fig. 1-(vi)).

With the aim of replicating the clinical protocol for fetal brain MRI, FSE acquisitions are simulated in the three orthogonal orientations. Besides, the position of the field-of-view is slightly shifted by ± 1.6 mm in the slice thickness orientation to produce additional partially-overlapping datasets in each orientation.

Fetal motion. The amplitude of typical fetal movements is estimated from clinical data^{71,72}. Three levels are defined accordingly for little, moderate and strong motion of the fetus, with a maximum of 5% corrupted slices over the fetal brain volume. They are characterized by a uniform distribution of respectively $[-1, 1]$ mm, $[-3, 3]$ mm and $[-4, 4]$ mm for independent translation in every direction and $[-2, 2]^\circ$, $[-5, 5]^\circ$ and $[-8, 8]^\circ$ for 3D rotation (Fig. 1-(v)).

Computational specifications. Since the addition of 3D motion during k-space sampling is expensive in computing memory, the simulations are run on 16 CPU workers in parallel with 20 GB of RAM each. After investigation, the amount of resources allocated to run such simulations can be reduced to 13 CPU workers in parallel with 20 GB of RAM each.

Simulating data from clinical MR acquisitions. Clinical datasets. The data used in this study were acquired in earlier studies in accordance with the relevant guidelines and regulations, under the supervision of Ethics Boards composed of representatives at different levels (hospitals, cantons, and federal state). Mothers of all fetuses included in the current work provided written informed consent for the re-use of their data for research purposes. Clinical cases are used as typical acquisition examples to generate realistic synthetic images of the fetal brain throughout development and to visually compare the quality of the simulated images.

Thirteen healthy subjects in the GA range of 21 to 33 weeks (27.0 ± 3.85 weeks) were scanned at CHUV as part of a larger institutional research protocol approved by the ethical committee of the Canton of Vaud, Switzerland (CER-VD, decision number: 2021-00124). In particular, these clinical cases are used to showcase the implemented pipeline and the realistic appearance of the corresponding synthetic HASTE images, as well as to explore the potential of FaBiAN in optimizing SR fetal brain MRI. For this purpose, clinical acquisitions were reconstructed offline using the docker version of the MIAL Super-Resolution Toolkit^{16,52}.

Fifteen subjects (thirteen neurotypical subjects and two subjects with light ventriculomegaly) from the Fetal Tissue Annotation Dataset (FeTA)²⁵ in the GA range of 21 to 34.6 weeks (27.5 ± 4.46 weeks) were scanned at Kispi. Their inclusion in research studies was approved by the ethical committee of the Canton of Zurich, Switzerland (KEK, decision number: 2016-01019). The original LR series of each subject were combined with each other *a posteriori* and offline to form a SR volume of the fetal brain using a simplified version of the Image Registration Toolkit (SIMPLE IRTK)¹⁴ under Licence from Ixico Ltd. The resulting SR reconstructions were manually segmented according to the FeTA annotation guidelines²⁵. In particular, the clinical cases from Kispi allow to extend FaBiAN to generate multiple SS-FSE images of the fetal brain with various settings. We investigate if these simulated images are realistic enough to replace part of the original data in the training phase of a deep learning network for fetal brain tissue segmentation, and to complement a clinical dataset to improve the performance of the segmentation algorithm.

Clinical MRI protocol. Typical fetal brain HASTE images are acquired on patients at 1.5 T (MAGNETOM Aera, Siemens Healthcare, Erlangen, Germany) with an 18-channel body coil and a 32-channel spine coil at CHUV. At

MR vendor	Magnetic field (T)	Clinical acquisitions		Simulated acquisitions	
		Number of subjects	GA (weeks)	Number of subjects	GA (weeks)
			Min–max (mean \pm SD)		Min–max (mean \pm SD)
Siemens Healthineers	1.5	13	21.0–33.0 (27.0 \pm 3.85)	10	21.0–33.0 (27.1 \pm 3.75)
GE Healthcare	1.5	6	21.0–34.6 (27.3 \pm 5.27)	6	21.0–35.0 (27.3 \pm 5.54)
	3	9	21.3–33.0 (27.6 \pm 4.18)	9	22.0–34.0 (28.4 \pm 3.91)

Table 3. Number of subjects, either scanned or simulated, considered throughout this study and distribution of gestational age (GA) according to the MR vendor and the main magnetic field strength.

least three T2w series (3–11 series, 6.69 ± 2.14 series) of 2D thick slices are acquired in three orthogonal orientations (axial, coronal and sagittal) with respect to the fetal brain using an ultra-fast multi-slice HASTE sequence (TR/TE, 1200 ms/90 ms; flip angle, 90° ; echo train length, 224; echo spacing, 4.08 ms; field-of-view, 360×360 mm²; voxel size, $1.13 \times 1.13 \times 3.00$ mm³; inter-slice gap, 10%)^{20,73}. Twenty-two to thirty slices are needed to cover the whole fetal brain depending on the GA (between 21 and 33 weeks) and size of the fetus, which corresponds to an acquisition time of between 26 to 36 seconds.

Fetal brain SS-FSE images are acquired on patients on either a 1.5-T or 3-T clinical GE whole-body scanner (Signa Discovery MR450 or MR750), either using an 8-channel cardiac coil or body coil, at Kispi. At least three series (4–6 series, 5.53 ± 0.64 series) of 2D thick slices are acquired in three orthogonal orientations (axial, coronal and sagittal) with respect to the fetal brain using a T2w SS-FSE sequence (TR/TE, 3000–3200 ms/116.032–124.08 ms; flip angle, 90° ; echo train length, 224; echo spacing, 3–3.2 ms; field-of-view, from 240×240 mm² to 300×300 mm²; slice thickness, 3.00–4.00 mm; acquisition matrix, 1.5 T: 256×224 voxels², 3 T: 320×224 voxels²; reconstruction matrix, 512×512 voxels²; isotropic in-plane resolution, from 0.47×0.47 mm² to 0.59×0.59 mm²)²⁵. Twenty-four to forty-three slices are needed to cover the whole fetal brain depending on the GA (between 21 and 34.6 weeks) and size of the fetus, which corresponds to an acquisition time of between one to two minutes.

The position of the field-of-view is slightly shifted in the slice thickness orientation to acquire additional data with some redundancy. In clinical practice, a total of six partially-overlapping LR series are commonly acquired in the three orthogonal orientations for subsequent SR reconstruction of the fetal brain.

Simulated datasets. The general framework presented in this paper makes it possible to simulate the clinical acquisition schemes described above for different MR vendors, at various magnetic field strengths, and with realistic SNR and amplitude of fetal movements.

The clinical cases from CHUV are used as representative examples of fetal brain HASTE acquisitions: the corresponding sequence parameters are replicated to simulate HASTE images of the fetal brain at various GA in the GA range of 21 to 33 weeks (27.8 ± 3.74 weeks). The amplitude of fetal movements in clinical acquisitions is assessed by an engineer expert in MR image analysis to ensure a similar level of motion in the simulated images. Besides, a 3D HR 1.1-mm-isotropic HASTE image of the fetal brain is simulated without noise or motion to serve as a reference for the quantitative evaluation of SR reconstructions from simulated LR 1.1-mm-in-plane HASTE images using the docker version of the MIAL Super-Resolution Toolkit^{16,52}.

SS-FSE images of the fetal brain are simulated for fifteen subjects in the GA range of 21 to 35 weeks (28.0 ± 4.47 weeks) at either 1.5 T or 3 T. We reproduce the same acquisition parameters and geometry as in the clinical dataset from Kispi. For every subject, three partially-overlapping series are simulated in each of the three orthogonal orientations, two with little motion and one with moderate motion.

Table 3 reports the main characteristics of the original clinical acquisitions and the simulated data according to the MR vendor and the main magnetic field strength.

Motion index. An engineer in biomedical imaging visually inspected all the clinical data retrospectively. 46% of the MR images from CHUV were rated as corrupted by little motion, 38% by moderate motion, and the remaining 16% by strong motion. In the case of MR images from Kispi, 53% were rated as corrupted by little motion, 35% by moderate motion, and the remaining 12% by strong motion. A motion index is defined to support the assessment of the level of fetal movements in clinical acquisitions and simulations based on binary masks that are drawn on each LR series to cover the whole fetal brain volume using ITK-SNAP⁷⁴. The motion index is estimated from tracking the displacement of the centroids of the 2D brain masks. It is computed as the sum of the variances of the 2D brain mask centroid coordinates of adjacent slices over the central third of the 3D brain mask, normalized by the number of slices considered (see Supplementary Method S2). An index less than 0.5 mm stands for little motion, in the range of [0.5,1]mm for moderate motion, and larger than 1 mm for strong motion. This motion index correlates well with the assessment of motion in multiple MR images from CHUV by an engineer expert in MR image analysis. It is used in the following to estimate the amplitude of fetal movements in clinical acquisitions and ensure a similar level of motion in the simulated images.

Qualitative assessment. A neuroradiologist and a pediatric (neuro)radiologist provided independent qualitative assessment of the HASTE images of the fetal brain simulated in the GA range of 21 to 33 weeks, in the three orthogonal orientations and with various levels of motion. Special attention was paid to the MR contrast between brain tissues, the SNR, the consistency of the brain anatomy, the delineation and sharpness of the structures of diagnostic interest (i.e., corpus callosum, vermis, brain stem, lateral ventricles, cortex, white matter, pituitary gland, etc.) that are analyzed in clinical routine, as well as characteristic motion artefacts. In a second step, six months later, the same neuroradiologist (Rater 1) and pediatric (neuro)radiologist (Rater 2) went through most of the simulated images, including SS-FSE images, and evaluated their realistic appearance and their similarity to MR images acquired in clinical routine based on a quality index between 0 and 2: 0 corresponds to data that do not look realistic, 1 to quite realistic data, meaning they have similarities with clinical acquisitions, and 2 to highly realistic data.

To facilitate visualization and qualitative comparison of images throughout this manuscript, simulated brain images were co-registered with clinical acquisitions of fetuses at the same GA and with equivalent level of motion. For this purpose, at least five landmarks were manually defined over the fetal brain volume using the landmark registration of 3D Slicer⁷⁵.

Application 1: Super-resolution reconstruction. *Implementation of SR reconstruction.* Orthogonal T2w LR HASTE series contain redundant information that enables the subsequent reconstruction of a 3D HR volume. Clinical acquisitions, respectively simulated images, are combined into a motion-free 3D image $\hat{\mathbf{X}}$ using the Total Variation (TV) SR reconstruction algorithm^{16,52} which solves:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \frac{\lambda}{2} \sum_{kl} \underbrace{\|\mathbf{D}_{kl}\mathbf{B}_{kl}\mathbf{M}_{kl}\mathbf{X} - \mathbf{X}_{kl}^{LR}\|^2}_{\mathbf{H}_{kl}} + \|\mathbf{X}\|_{TV}, \quad (1)$$

where the first term relates to data fidelity with k being the k -th LR series \mathbf{X}^{LR} and l the l -th slice, $\|\mathbf{X}\|_{TV}$ is a TV prior introduced to regularize the solution, and λ balances the trade-off between data fidelity and regularization terms (default setting, $\lambda = 0.75$). \mathbf{D} and \mathbf{B} are linear downsampling and Gaussian blurring operators given by the acquisition characteristics. \mathbf{M} encodes the rigid motion of slices.

Regularization setting. LR HASTE images of the fetal brain are simulated to mimic clinical HASTE acquisitions of three subjects of 26, 30 and 33 weeks of GA respectively, which are characterized by a high SNR, the presence of a realistic low to moderate bias field, and variable amplitudes of fetal movements, from little to strong. Particular attention is paid to ensuring that the motion level is respected. For each subject, a SR volume of the fetal brain is reconstructed from the various orthogonal acquisitions, either real or simulated (six, respectively seven and eight series were acquired and are further simulated in the subject of 26, respectively 30 and 33 weeks of GA), with different values of λ (0.1, 0.3, 0.5, 0.75, 1.5, 3) to study the potential of FaBiAN in optimizing the quality of the SR reconstruction in a clinical setup. Indeed, the weight λ is a sensitive hyper-parameter in SR methods that are based on solving an inverse problem. A quantitative analysis is conducted on the resulting SR reconstructions to determine the value of λ that provides the sharpest reconstruction of the fetal brain with high SNR, namely the smallest normalized root mean squared error (NRMSE) with respect to the corresponding synthetic 3D HR ground truth.

Number of LR series: an SNR and motion case study. Static LR HASTE images of the fetal brain (GA of 30 weeks) are recovered after the addition of various levels of complex Gaussian noise (mean, 0; standard deviation, 0.07, 0.15 or 0.3) to k-space. A standard deviation (SD) of 0.15 results in synthetic images that closely resemble clinical acquisitions. Six independent realizations of the 2D series are generated for each SNR. A SR volume of the fetal brain is reconstructed from three, six and nine orthogonal LR HASTE series using the different realizations. In total, fifty-four SR volumes are reconstructed.

Additional LR HASTE acquisitions of the fetal brain (GA: 30 weeks, SD of noise: 0.15) are simulated with inter-slice motion. The impact of little and moderate movements of the fetus on SR reconstructions from various numbers of LR series is studied using a reference series without motion and with little amplitude of fetal motion respectively.

The different configurations are compared to each other by computing the NRMSE, the local SSIM and its mean (MSSIM) over the image with respect to the 3D isotropic ground truth. The latter are computed using Matlab *ssim* function with the SD of the isotropic Gaussian function that determines the weights of the pixels in a neighborhood to estimate local statistics set to its default value (1.5), and the dynamic range of the normalized input images to 255.

Application 2: Data augmentation for automated fetal brain tissue segmentation. Supplementary Figure S2 displays the histograms of the distribution of (a) gestational age (from 21 to 35 weeks, 27.8 ± 4.40 weeks) across the original clinical cases and simulated subjects involved in this experiment, as well as (b) in-plane isotropic resolution in the corresponding LR images according to the main magnetic field strength.

LR SS-FSE images of the fetal brain are simulated based on the MR sequence parameters and acquisition settings extracted from the clinical cases scanned at Kispi to mimic typical MR acquisitions, thus minimizing confounding factors. The synthetic images are interpolated to $0.8594 \text{ mm} \times 0.8594 \text{ mm}$ in the in-plane direction to match the resolution of the clinical SR reconstructions. Label maps are automatically generated from the simulation framework using a nearest-neighbour interpolation. Two different labels are assigned to the

	Configuration	Number of clinical subjects	Number of simulated subjects	Total number of subjects	Number of replicates
Experiment 1	(A)	15	0	15	2
	(B)	10	5	15	2
	(C)	8	7	15	2
Experiment 2	(D)	15	15	30	2
	(E)	15	0	15	4

Table 4. Two experiments are presented along with the different configurations studied to compare the performance of an algorithm for automated fetal brain tissue segmentation. For each configuration, the respective number of clinical cases and simulated subjects, and the total number of subjects involved in the cross-validation are reported, as well as the number of times standard augmentation is performed (number of replicates).

ventricular system (lateral, third and fourth ventricles) and to the extra-axial cerebrospinal fluid spaces in the SR reconstructions from Kispi subjects²⁵, whereas the segmented HR anatomical images from which our simulations are derived distinguish the lateral ventricles from the cerebrospinal fluid¹⁰. For consistency between brain annotations, we merge the labels that correspond to structures from the ventricular system and the cerebrospinal fluid spaces in both models.

Network architecture. We designed a convolutional neural network based on the well established U-Net architecture⁷⁶ for biomedical semantic image segmentation, as it recently proved its ability to perform well for 2D fetal brain MRI tissue segmentation²². The baseline 2D U-Net is trained using a hybrid loss function defined as the sum of a categorical cross-entropy and a Dice loss. The latter is intended to mitigate any imbalance in the samples of the different classes^{22,77}.

The implementation is performed in the framework of TensorFlow 2.5⁷⁸ and an Nvidia GeForce RTX 2080 GPU is deployed for training.

Experimental design. As indicated in Table 4, we run two independent experiments:

- **Experiment 1:** It aims at investigating whether the simulated images are realistic enough to substitute for clinical data. We compare three configurations that combine different proportions of subjects with SR volumes from clinical MR acquisitions and subjects with simulated LR images: a baseline (A) that consists of real data only (fifteen subjects), a configuration (B) that gathers ten original subjects and five simulated ones, a configuration (C) with eight original subjects and seven simulated ones. Real data remain predominant in all configurations studied.
- **Experiment 2:** If the MR images generated using FaBiAN are close enough to real data, we hypothesize that they could complement a small clinical dataset. In this setting, synthetic images are used as additional data to augment the training dataset. The configuration (D) complements the fifteen original subjects from the baseline (A) with fifteen simulated ones. It is compared to a conventional data augmentation strategy (E) that includes only the fifteen original subjects from the baseline (A).

Since we could not automatically propagate the annotations on the simulated images to the corresponding SR reconstruction, and conversely the labels in the SR reconstruction from clinical acquisitions to the original LR 2D series, three LR series are needed to replace one SR volume. The segmentation is computed on skull-stripped images to process only the voxels within the intracranial volume.

Training strategy. Networks are fed with 64-by-64 overlapping patches in the axial orientation. To avoid any bias, we ensure an equivalent number of 2D patches between an SR-reconstructed clinical case and three LR axial series simulated for a given subject. Intensities of all image patches are standardized.

- **Experiment 1:** Each patch is repeated once using minimal data augmentation, i.e., random flip and rotation of the patches (n times, by $n \times 90^\circ$, $n \in \llbracket 0, 3 \rrbracket$).
- **Experiment 2:** To maintain a similar number of input samples, each patch from configuration (D) is repeated once, while patches from configuration (E) are repeated three times. (D) is only augmented through random flip and rotation, while more extensive conventional augmentation strategies used in fetal brain MRI segmentation²⁵, which also include random Gaussian noise, random gamma, and random bias field, are applied in configuration (E)⁷⁹.

Five-fold cross-validation approaches (training sets: 12 subjects, validation sets: 3 subjects) are used to determine the epochs for the learning rate decay in each configuration.

Analysis. The performance of the fetal brain tissue segmentation networks is evaluated based on the Dice similarity coefficient (DSC)⁸⁰ that quantifies the overlap between the predicted segmentation and the manually-annotated ground truth. The performance metrics are assessed on the validation sets. We report the average over all folds. Statistical significance is given between both data augmentation strategies (configurations (D) and (E)) in Experiment 2. *P*-values of Wilcoxon rank sum test for individual fetal brain tissue segmentation are adjusted for multiple comparisons using Bonferroni correction. *P* < 0.05 is considered statistically significant.

Results

Computational performance. As highlighted in Fig. 1 for a fetus of 30 weeks of GA whose brain is covered by twenty-five slices, the computation time to convert segmented HR anatomical images of the fetal brain to MR contrast is in the order of one second. EPG simulations are run in every voxel of the 3D HR anatomical images in less than four minutes. For one axial series with either little, moderate or strong motion, k-space sampling takes less than seven minutes for HASTE images simulated with an acceleration factor of two, respectively less than eight minutes for SS-FSE images simulated without using any acceleration technique.

Qualitative assessment. Figure 2 illustrates the close resemblance between simulated HASTE images of the fetal brain and clinical MR acquisitions in terms of MR contrast between tissues, SNR, brain anatomy and relative proportions across development for representative subjects in the GA range of 23 to 32 weeks, as well as typical out-of-plane motion patterns related to the interleaved slice acquisition scheme. A neuroradiologist and a pediatric (neuro)radiologist report a good contrast between gray and white matter, which is important to investigate the cortex continuity and identify the deep gray nuclei as well as any neuronal migration defect. The radiologists also notice good SNR in the different series and report proper visualization of the main anatomical structures: the four ventricles, the corpus callosum, the vermis, the cerebellum, even sometimes the fornix. Besides, the experts are able to monitor the evolution of normal gyration throughout gestation. However, the radiologists point out that small structures such as the pituitary gland, the chiasma, the recesses of the third ventricle, and the vermis folds that look part of the cerebellum, are more difficult to observe. The cortical ribbon is clearly visible but quite pixelated, which is likely to complicate the diagnosis of polymicrogyria. White matter appears too homogeneous, which makes its multilayer aspect barely distinguishable, with an MR signal that is constant across GA, thus preventing physicians from exploring the myelination process throughout brain maturation.

Further investigation is illustrated in Table 5 based on a quality index, where Rater 1 refers to the neuroradiologist and Rater 2 to the pediatric (neuro)radiologist. Most images are evaluated as highly realistic, or at least quite realistic: 97%, respectively 86% of the HASTE images are evaluated with an index greater than or equal to 1 by Rater 1, respectively Rater 2, and 92%, respectively 78% of the SS-FSE images are evaluated with an index greater than or equal to 1 by Rater 1, respectively Rater 2. The better evaluation of the simulated HASTE images may be related to the fact that the raters are used to analyzing HASTE acquisitions in their daily practice rather than SS-FSE scans. Rater 1 emphasizes that synthetic images look even more realistic in older fetuses. Rater 1 also highlights that FaBiAN reproduces incredibly well fetal movements as they occur in clinical practice. It is worth noting that this index is intended to better reflect the *realism* of the simulated images rather than the actual *image quality*. However, especially for those corrupted by high levels of fetal motion, the low quality of the resulting images may impact the overall appreciation of the raters. Indeed, the HASTE images that have been evaluated with an index of 0 are, in most cases (> 83%), corrupted by a moderate to strong level of motion. More than half of the SS-FSE images that have been evaluated as not realistic are simulated in the sagittal orientation. Rater 2 points out that sagittal views of fetuses of GA between 32 and 35 weeks show cerebrospinal fluid above the posterior part of the corpus callosum, which is not as pronounced in clinical images. In addition, Rater 2 explains that images with significant duplication of a structure are also rated at 0 because such artefacts do not appear as prominent in clinical practice, but either the image is not in a strict orientation or part of it is darker due to signal drops.

Application 1: Super-resolution reconstruction. *Regularization setting.* Thanks to its controlled environment, FaBiAN makes it possible to adjust the parameter λ for optimal SR reconstruction with respect to a simulated 3D isotropic HR ground truth of the fetal brain. Figure 3 explores the quality of SR fetal brain MRI from LR HASTE images corrupted by motion depending on the weight of TV regularization in two subjects of 26 and 30 weeks of GA respectively. Based on the simulations, a high level of regularization ($\lambda = 0.1$) provides a blurry SR reconstruction with poor contrast between the various structures of the fetal brain, especially in the deep gray nuclei and the cortical plate. In addition, the cerebrospinal fluid appears brighter than in the reference image. A low level of regularization ($\lambda = 3$) leads to a better tissue contrast but increases the overall amount of noise in the resulting SR reconstruction. A fine-tuned regularization ($\lambda = 0.75$) provides a sharp reconstruction of the fetal brain with a high SNR and a tissue contrast close to the one displayed in the reference image. In the SR images reconstructed from clinical LR HASTE series altered by a little-to-moderate level of motion, as in the simulations, the structure of the corpus callosum and the delineation of the cortex are especially well defined for appropriate TV regularization ($\lambda = 0.75$), leading to high-SNR HR images of the fetal brain. Although the NRMSE between SR reconstructions from simulated HASTE images and the corresponding ground truth are close to each other for a given GA across the various weights studied, Fig. 4 shows that the error is systematically minimal for $\lambda = 0.75$, which further supports this parameter setting for optimal SR reconstruction of the fetal brain from this type of MR images.

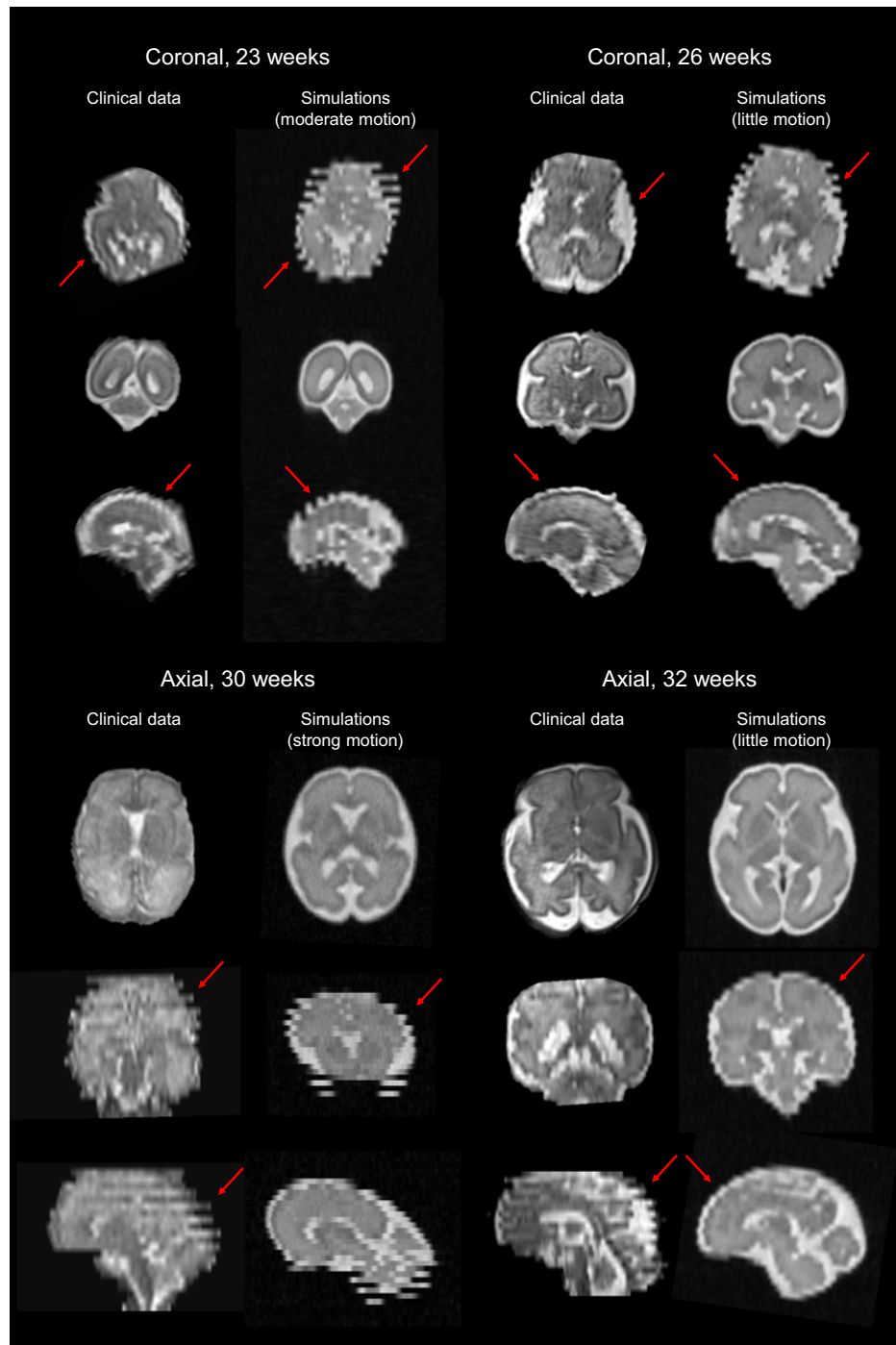


Figure 2. Visual inspection and comparison between clinical MR acquisitions and representative simulated HASTE images of the fetal brain in the three orthogonal orientations at four different GA (23, 26, 30 and 32 weeks). The amplitude of movement of the fetus is indicated from the motion index computation. Red arrows point out typical out-of-plane motion patterns.

Number of LR series: an SNR and motion case study. Figure 5 shows the NRMSE and the MSSIM between SR reconstructions from different numbers of orthogonal LR HASTE series simulated with various SNR or variable amplitude of movements and a 3D HR reference⁸¹.

In the case of static data (Fig. 5-left panel), the NRMSE decreases when increasing the number of series used in the SR reconstruction. According to both the NRMSE and the MSSIM, the quality of the SR reconstructions resulting from simulated images with an SNR close to that observed in clinical acquisitions and from synthetic images with a distribution of complex Gaussian noise of twice less SD is similar. Noisier images lead to a slight decrease in the NRMSE, but also in the MSSIM which in turn increases with the number of series.

Index	Simulated HASTE images			Simulated SS-FSE images		
	2	1	0	2	1	0
Rater 1	54	43	3	56	36	8
Rater 2	74	12	14	52	26	22

Table 5. Independent evaluation of the realism of the images generated using FaBiAN based on a quality index. Percentage of 2—highly realistic, 1—quite realistic, 0—non-realistic synthetic HASTE and SS-FSE images according to a neuroradiologist (Rater 1) and a pediatric (neuro)radiologist (Rater 2).

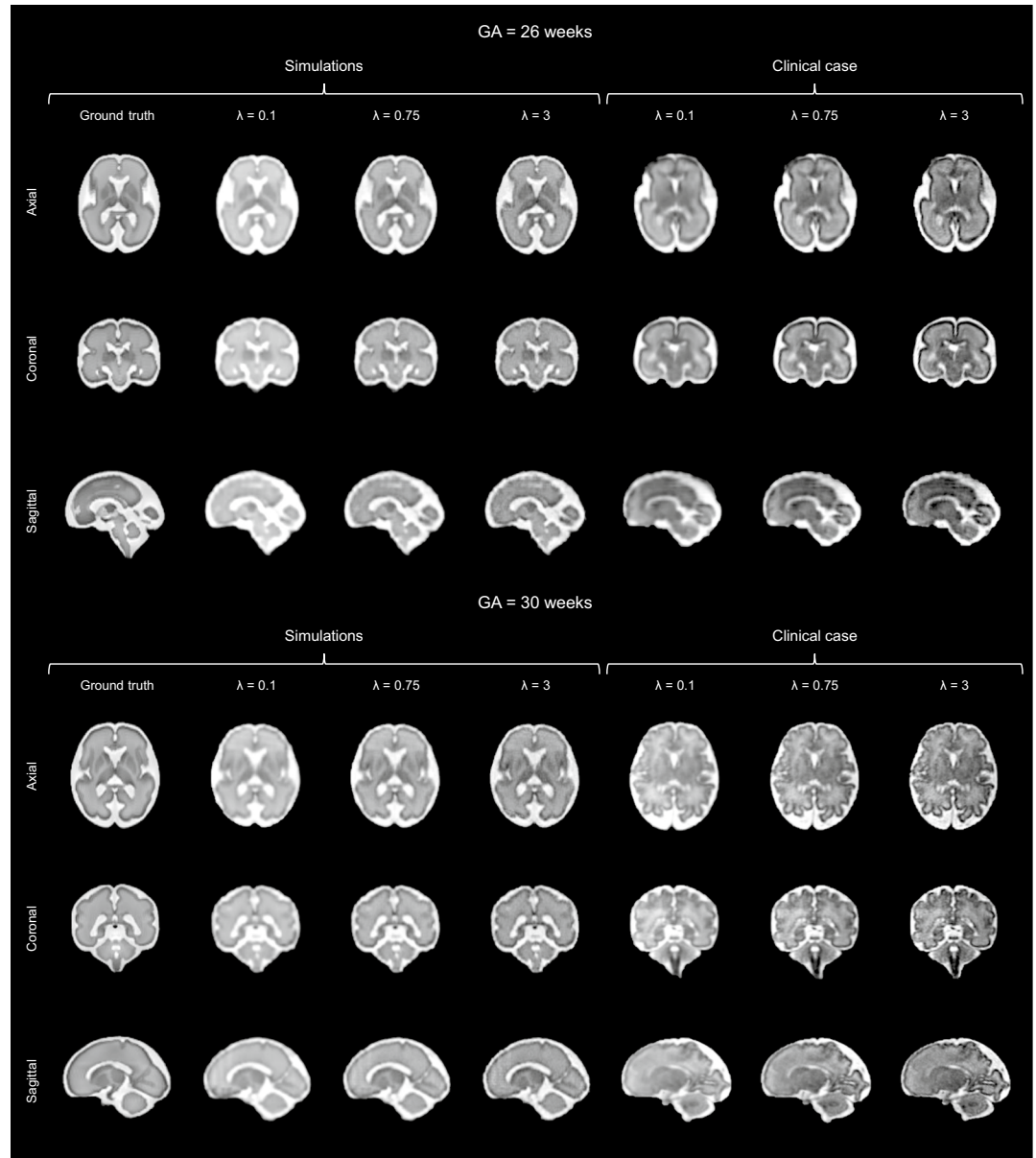


Figure 3. Appreciation of the quality of SR reconstruction depending on the weight λ that controls the strength of the TV regularization. The potential of our framework FaBiAN for optimizing the reconstruction quality through parameter fine-tuning in the presence of motion is illustrated at two GA: 26 and 30 weeks. Two representative clinical cases are provided for comparison. The results for three values of λ are presented. For $\lambda = 0.1$, the SR reconstruction looks blurry with poor tissue contrast. Using $\lambda = 3$ improves the contrast but the images look noisy. For $\lambda = 0.75$, the SR reconstruction is sharp with a contrast between different brain tissues similar to that observed in the 3D isotropic ground truth. Clinical cases from which the simulated HASTE images are derived highlight the accuracy of a SR reconstruction for this intermediate value of λ , especially with regards to the definition of the corpus callosum and the delineation of the cortex.

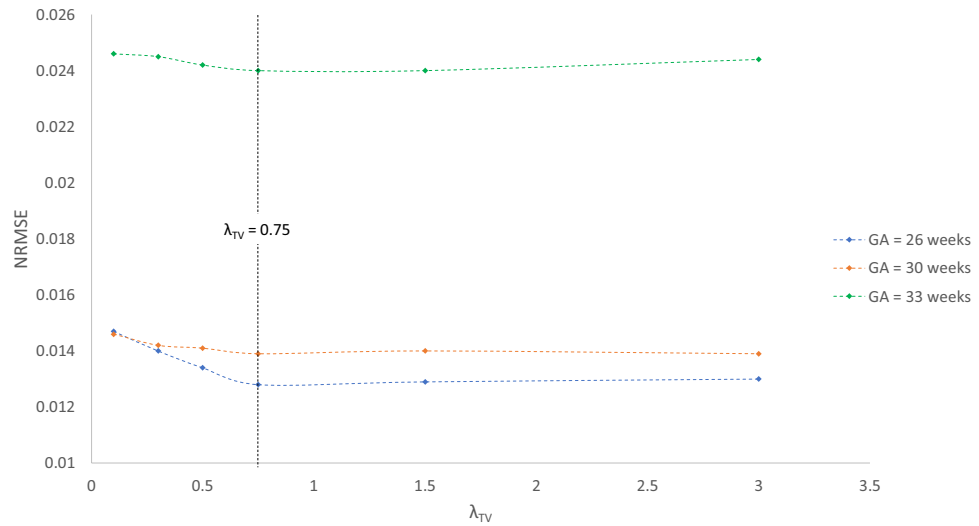


Figure 4. Normalized root mean squared error (NRMSE) between SR reconstructions from simulated data at a GA of 26, 30 and 33 weeks respectively and the corresponding 3D HR ground truth depending on the weight λ of the TV regularization. Six values of λ are tested: 0.1, 0.3, 0.5, 0.75, 1.5 and 3. The NRMSE is minimal for $\lambda = 0.75$.

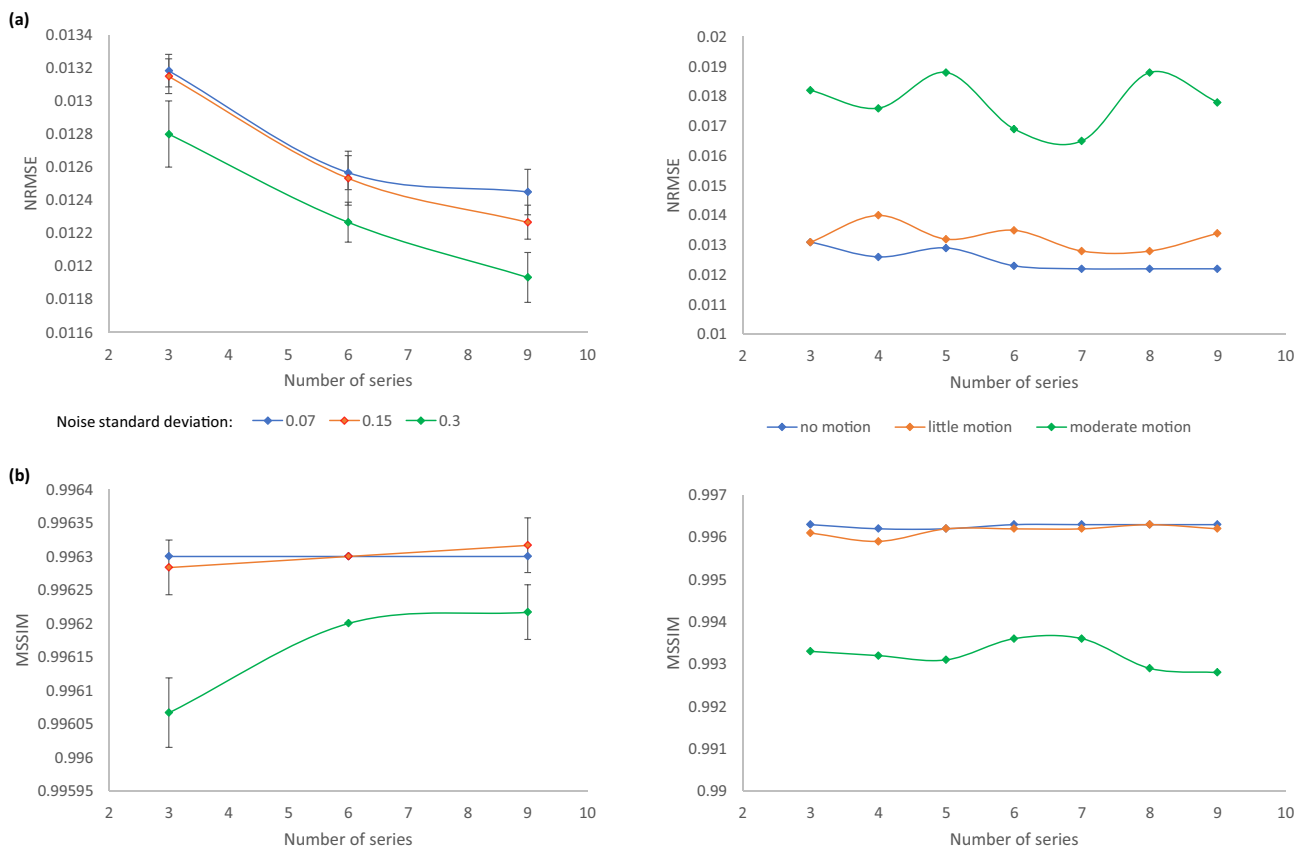


Figure 5. (a) Normalized root mean squared error (NRMSE) and (b) mean structural similarity index (MSSIM) between SR reconstructions from different numbers of orthogonal LR HASTE series simulated at a GA of 30 weeks and the corresponding static 3D HR ground truth. The left panel shows results for motion-free data with various noise levels, a SD of 0.15 leading to a similar appearance as in clinical acquisitions. The right panel illustrates how the algorithm performs depending on the amplitude of fetal movements in the input series.

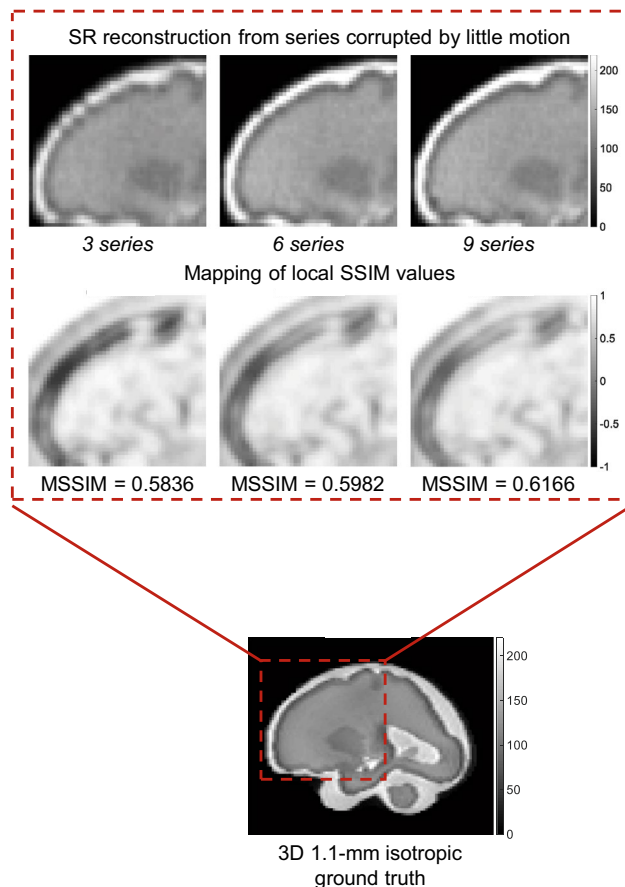


Figure 6. Appreciation of sharpness and tissue contrast enhancement in SR reconstructions from higher numbers of simulated orthogonal LR HASTE images corrupted by little motion at a GA of 30 weeks in comparison with the corresponding static 3D HR ground truth. The frontal cortex looks smoother and the putamen area sharper in the SR reconstruction from nine series compared to the SR reconstruction from three series. The mapping of local SSIM values and the computation of the MSSIM over the corresponding region-of-interest further support these observations.

The stronger the movements in the LR series, the higher the NRMSE of the resulting SR reconstruction. The addition of motion-corrupted LR series to reconstruct a SR volume of the fetal brain does not increase the MSSIM. Data with slight motion are well handled by the SR algorithm as the MSSIM is equivalent for SR reconstructions from static images and series with little motion. In the case of moderate motion, the MSSIM is lower than in the case of little motion.

Figure 6 highlights the benefit of increasing the number of orthogonal LR series on the rendering of the SR reconstruction. The higher the number of LR series combined in the SR reconstruction, even altered by motion, the smoother the frontal cortex and the sharper the putamen area in the resulting SR volume⁸¹.

Application 2: Data augmentation for automated fetal brain tissue segmentation. We aim at proving the realistic appearance of the simulated images (representative SS-FSE images of two subjects scanned/simulated at 1.5 T and 3 T respectively are provided as Supplementary Fig. S3) and demonstrating the practical value of the developed environment to complement clinical datasets for data augmentation strategies in deep learning, here with the example of fetal brain tissue segmentation.

Table 6 shows the mean DSC \pm SD computed for every segmented brain tissue in each configuration. Overall, the performance of the segmentation algorithm is maintained when replacing original subjects by synthetic images obtained from an SS-FSE sequence simulated with the same acquisition parameters as in the clinical protocol (Table 6, Experiment 1). This trend is also observed for each individual structure studied. In the configuration (C), the overall DSC for brain tissue segmentation is slightly increased compared to the baseline.

Since the simulated images look realistic enough to substitute for original clinical acquisitions, we further investigate if they can be used for data augmentation (Table 6, Experiment 2). Increasing the training data (15 real cases) by 15 supplementary simulated subjects (configuration D) results in a significantly improved mean DSC of 0.90 ± 0.05 over the six segmented brain tissues compared to 0.86 ± 0.06 in configuration (E), where the 15 original cases are extensively augmented by conventional intensity-based operations such as additional Gaussian noise, bias field and gamma intensity changes⁷⁹. In more detail, the DSC is higher for all segmented brain structures when complementing clinical acquisitions with close simulated data, with statistical significance

	Experiment 1			Experiment 2	
	(A) C15/S0	(B) C10/S5	(C) C8/S7	(D) C15/S15	(E) C15/S0
CSF & ventricles	0.93 ± 0.01	0.94 ± 0.01	0.94 ± 0.02	0.95 ± 0.02 (*)	0.93 ± 0.01
Cortical GM	0.77 ± 0.02	0.80 ± 0.04	0.81 ± 0.05	0.84 ± 0.05 (*)	0.77 ± 0.02
WM	0.92 ± 0.01	0.92 ± 0.02	0.92 ± 0.02	0.93 ± 0.02	0.92 ± 0.01
Cerebellum	0.88 ± 0.04	0.87 ± 0.10	0.87 ± 0.09	0.92 ± 0.04 (*)	0.87 ± 0.06
Deep GM	0.85 ± 0.03	0.84 ± 0.10	0.87 ± 0.04	0.90 ± 0.04 (*)	0.85 ± 0.04
Brain stem	0.84 ± 0.03	0.85 ± 0.04	0.86 ± 0.04	0.88 ± 0.03 (*)	0.85 ± 0.03
Overall	0.87 ± 0.06	0.87 ± 0.08	0.88 ± 0.06	0.90 ± 0.05 (*)	0.86 ± 0.06

Table 6. DSC (mean ± SD) in the different configurations studied for all segmented brain tissues: cerebrospinal fluid (CSF) and ventricles, cortical gray matter (GM), white matter (WM), cerebellum, deep gray matter and brain stem, and on average. The number of clinical cases (Cxx) and the number of simulated subjects (Sxx) are recalled. We also emphasize that augmentation is performed twice more in configuration (E) compared to other configurations. The segmentation algorithm performs better (score in bold) in every structure when complementing the baseline dataset (configuration (A)) with simulated subjects (configuration (D)) than when performing standard data augmentation (configuration (E)). *P*-values of Wilcoxon rank sum test between both data augmentation strategies (configurations (D) and (E)) for individual fetal brain tissue segmentation are adjusted for multiple comparisons using Bonferroni correction. *P* < 0.05 (*) is considered statistically significant.

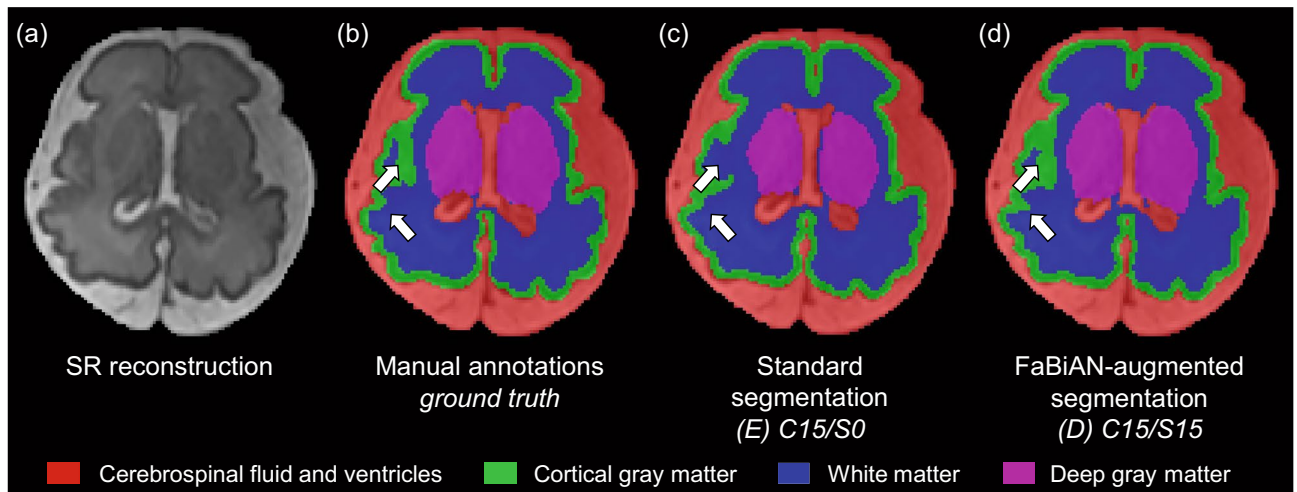


Figure 7. Illustration of the accuracy of fetal brain tissue segmentation in a subject of 30.6 weeks of GA on (a) an axial slice from the SR reconstruction. Comparison of (b) the reference manual annotations, (c) the segmentation results obtained when performing extensive standard data augmentation on the clinical SR reconstructions (configuration (E), C15/S0), (d) the segmentation results obtained by the configuration (D) that complements this original dataset with fifteen additional simulated subjects (C15/S15), overlaid on the SR image. The segmentation of the cortex especially looks more accurate in (d), with an increased sensitivity to folding as highlighted by the white arrows.

(*p* < 0.05) for the cerebrospinal fluid and ventricles, the cortical gray matter, the cerebellum, the deep gray matter and the brain stem.

Figure 7 illustrates on an axial view the accuracy of fetal brain tissue segmentation in a subject of 30.6 weeks of GA. The results obtained in the configuration (D) where real clinical data are complemented with simulated subjects look close to the manually-annotated ground truth. In particular, the segmentation seems to be more accurate in the cortex with an enhanced sensitivity to folding compared to the segmentation obtained in the baseline (A) where the network is solely trained on clinical data.

Discussion and conclusion

In this work, we present FaBiAN, a novel Fetal Brain magnetic resonance Acquisition Numerical phantom, and illustrate some of its potential uses. Our tool relies on EPG simulations to account for stimulated echoes in the computation of the T2 decay in every voxel of the HR anatomical images from which the simulated images are derived. The developed framework remains general and highly flexible in the choice of the sequence parameters and anatomical settings available to the user. It simulates as closely as possible the physical principles involved in

FSE sequences on several MR systems, namely from various MR vendors and at different magnetic field strengths, resulting in highly realistic T2w images of the developing brain throughout gestation.

The limitations in the resemblance of the simulated FSE images compared to typical clinical MR acquisitions may be explained by the origin of the simulated images and the lack of T1 and T2 ground truth measurements, both in the multiple fetal brain tissues and throughout maturation. As a model for our simulations, we use a normative spatiotemporal MRI atlas of the fetal brain¹⁰ where representative images at each GA correspond to an average of fetal brain scans across several subjects, thus resulting in smoothing of subtle inter-individual heterogeneities, especially in the multilayer aspect of the white matter. T1 and T2 mapping of the fetal brain is currently limited by long scanning times, unpredictable fetal motion, and acquisition of thick slices to ensure a good SNR over the whole fetal brain volume. For ethical reasons, there is also a lack of normative T1 and T2 values across GA. Therefore, and as a first approximation, we consider average T1 and T2 values of the various fetal brain structures labeled as gray matter, white matter or cerebrospinal fluid throughout development. As a result, our simulated images may fail to capture the fine details of the fetal brain anatomy as well as maturation processes that imply changes in T1 and T2 relaxation times during gestation. In this sense, experts report that they would feel confident in performing standard biometric measurements on the simulated images and in evaluating the volume of white matter, but not its fine structure. It is worth noticing that in-plane motion artefacts like signal drops are not accounted for at this stage, as slices severely corrupted by such artefacts will be removed from the analysis. Furthermore, we restricted our simulations to images of the fetal brain without the remaining fetal anatomy or the surrounding maternal uterine cavity, as most image analysis and post-processing techniques intended for fetal brain MRI, including SR reconstruction, first perform an automated fetal brain extraction/segmentation. Unlike other models that represent fetal subjects by scaled pediatric data and oversimplify the complexity of the fetal brain anatomy by reducing it to white matter alone or to a single brain tissue label^{82–84}, we wanted our numerical simulations to be based on a comprehensive model of the fetal brain consistent with the underlying clinical application. In this sense, the normative spatiotemporal MRI atlas of the fetal brain¹⁰ from which the simulated FSE images are built captures the details of the fetal brain anatomy throughout maturation. Future work will improve the accuracy of the fetal brain model, both in terms of structure differentiation and T1 and T2 changes during maturation. This implies modeling of T1 and T2 variations, both locally and across time, especially within white matter. We will generate locally varying reference relaxometric properties assigned to every brain tissue to capture changes within finer structures, and modulate these values according to the gestational age of the fetus to enhance the MR contrast between the various brain tissues throughout gestation and simulate more realistic images of the developing fetal brain. Thanks to the flexibility of FaBiAN, the surrounding fetal and maternal anatomy, as well as any other structures, could easily be included as long as we have access to segmented HR images, for instance by resorting to more complex anatomical models such as the XCAT phantoms^{85–87}. Several models can also be combined.

We have designed this open-source simulator to aid in the development and validation of advanced image processing techniques dedicated to improving the analysis of fetal brain MR images and support accurate diagnosis. Despite SR reconstruction has already demonstrated its potential for accurate biometric measurements in the fetal brain^{18–20}, some parameters still need to be adjusted to the nature of the input LR images to provide optimal evaluation and support computer-assisted diagnosis. In fetal MRI, the level of regularization is commonly set empirically based on visual perception^{12–14,17}. Intuitively, the level of regularization depends on the amount of data available to solve the ill-posed inverse problem. Thanks to its controlled environment, the presented framework makes it possible to explore the optimal settings for SR fetal brain MRI according to the quality of the input motion-corrupted LR series with respect to a synthetic 3D HR ground truth. It is worth noticing that in-plane motion artefacts like signal drops are not accounted for in the simulation workflow at this stage, as heavily corrupted slices are commonly removed from the reconstruction. Besides, FaBiAN also enables quantitative assessment of the robustness of any SR reconstruction algorithm depending on various parameters that can be intrinsic to the system like noise, or related to the clinical application such as the amplitude of fetal motion in the womb and the number of series used for SR reconstruction⁸¹. For instance, a decrease in the MSSIM when adding motion-corrupted series to reconstruct a SR volume of the fetal brain compared to the SR reconstruction from static series can be caused by an inappropriate slice-to-volume registration (SVR). Therefore, our numerical phantom provides a valuable framework for reproducibility studies and validation of image processing methods. Additional examples among its wide variety of applications include the simulation of a static reference volume at a given GA on which to align the clinical orthogonal LR series acquired in a subject at the same GA in order to perform SVR and subsequent SR reconstruction of the fetal brain, especially in the presence of heavily motion-corrupted acquisitions. Besides, the performance of the SVR can be quantitatively assessed by comparing the motion transform estimated by the algorithm to the controlled 3D rigid movements actually simulated in the images. Synthetic HR images can also be used as a reference for more general motion compensation techniques.

As raised by Wissmann et al., the lack of comparability between simulation setups hinders the evaluation of image reconstruction methods in relation to each other²⁷. This first numerical phantom for MR imaging of the in utero developing brain has already proven its high flexibility in generating multiple images with various acquisition parameters and settings. To make it even more general, we will extend FaBiAN to other MR vendors and sequences. From image acquisition to post-processing developments, the fetal brain MRI community should be able to take advantage of such a unified environment that is intended to simulate MR images of the developing fetal brain as they are acquired in clinical routine all around the world.

Furthermore, the developed framework generates T2w images of the fetal brain realistic enough to complement real clinical acquisitions for data augmentation strategies, as shown here with a proof of concept for fetal brain tissue segmentation. It especially makes it possible to exploit the whole range of GA in the simulated data when clinical cases are scarce and not necessarily uniformly distributed across development. Thus, we can take advantage of larger and more diverse datasets at no cost. In Experiment 1, we limit augmentation of the clinical

and simulated data to random flip and rotation only in order to minimize its effect in evaluating the combination of both types of images. Thus, it becomes possible to explore how replacing a real image with a synthetic one impacts the variability of the model, and subsequently the segmentation accuracy. In the scenario of Experiment 2, we rather compare standard data augmentation practices in modern convolutional neural network segmentation pipelines (additional Gaussian noise, bias field, and gamma intensity corrections)^{25,79} to the data augmentation possibilities offered by FaBiAN. Results show that the greater anatomical variability brought by the addition of MR images simulated from another source (configuration (D)) certainly benefits the segmentation accuracy more than intensity-based operations (configuration (E)). Overall, the SR images from the FeTA dataset used in this study present a homogeneous appearance, with high quality, denoised and bias-free reconstructions. Thus, augmentation of the original clinical dataset through noise addition and intensity changes does not help the network to better generalize in this cohort. Conversely, the segmentation network is more likely to take advantage of the more heterogeneous and diverse data generated by FaBiAN. In this proof-of-concept study, it is worth highlighting that the bias field and SNR are the same throughout all simulations. However, as demonstrated in this manuscript, FaBiAN is an excellent playground to perform additional data augmentation itself by generating multiple MR images for the same subject with varying parameters, whether related to intensity changes, spatial resolution, or amplitude of fetal movements. This asset would further complement conventional data augmentation strategies to enhance data heterogeneity and mitigate the scarcity of fetal brain MR images.

Although the resort to an atlas restricts the inter-subject variability at a given GA, the great flexibility of FaBiAN also lies in the possibility of simulating images from various sources, either atlases or clinical segmented HR anatomical images of the fetal brain like SR reconstructions. Thus, it becomes possible to simulate several subjects at a given GA from various fetal brain models to increase the inter-subject anatomical heterogeneity in the synthetic images. Beyond normal brain development, we aim at exploring common developmental pathologies, like ventriculomegaly and spina bifida, by simulating new datasets of synthetic MR images from annotated SR reconstructions of the fetal brain in pathological subjects.

In line with the demonstration of the added value of FaBiAN for diverse applications that revolve around improving diagnosis and prognosis from MR images of the developing fetal brain, future work involves investigating the ability of such a numerical framework to generalize post-processing tools like fetal brain tissue segmentation to datasets acquired on other MR systems and with other sequence parameters using a collection of various synthetic images for domain adaptation techniques.

Data availability

All the simulated datasets that support the various proof-of-concept studies documented in this manuscript will be made publicly available⁸⁸ upon acceptance of the present work.

Code availability

The proposed framework⁸⁹ is open source.

Received: 13 August 2021; Accepted: 5 April 2022

Published online: 23 May 2022

References

- Volpe, J. J. Brain injury in premature infants: A complex amalgam of destructive and developmental disturbances. *Lancet Neurol.* **8**, 110–124. [https://doi.org/10.1016/S1474-4422\(08\)70294-1](https://doi.org/10.1016/S1474-4422(08)70294-1) (2009).
- Ramirez-Vélez, R. In utero fetal programming and its impact on health in adulthood. *Endocrinol. Nutr. (English Ed.)* **59**, 383–393. <https://doi.org/10.1016/j.endonu.2012.02.002> (2012).
- Bhat, S., Acharya, U. R., Adeli, H., Bairy, G. M. & Adeli, A. Autism: Cause factors, early diagnosis and therapies. *Rev. Neurosci.* **25**, 841–850. <https://doi.org/10.1515/revneuro-2014-0056> (2014).
- Miller, S. L. *et al.* Antenatal antioxidant treatment with melatonin to decrease newborn neurodevelopmental deficits and brain injury caused by fetal growth restriction. *J. Pineal Res.* **56**, 283–294. <https://doi.org/10.1111/jpi.12121> (2014).
- Kwon, E. J. & Kim, Y. J. What is fetal programming?: A lifetime health is under the control of in utero health. *Obstet. Gynecol. Sci.* **60**, 506–519. <https://doi.org/10.5468/ogs.2017.60.6.506> (2017).
- O'Donnell, K. J. & Meaney, M. J. Fetal origins of mental health: The developmental origins of health and disease hypothesis. *Am. J. Psychiatry* **174**, 319–328. <https://doi.org/10.1176/appi.ajp.2016.16020138> (2017).
- Tiemeier, H. A closer look at the fetal programming hypothesis with obstetric ultrasound. *J. de Pédiatrie* **93**, 437–438. <https://doi.org/10.1016/j.jpmed.2017.04.001> (2017).
- Bilder, D. A. *et al.* Early second trimester maternal serum steroid-related biomarkers associated with autism spectrum disorder. *J. Autism Dev. Disord.* **49**, 4572–4583. <https://doi.org/10.1007/s10803-019-04162-2> (2019).
- Hayat, T. T. A. & Sami, M. B. Early magnetic resonance imaging biomarkers of schizophrenia spectrum disorders: Toward a fetal imaging perspective. *Dev. Psychopathol.* **33**, 1–15. <https://doi.org/10.1017/S0954579420000218> (2020).
- Gholipour, A. *et al.* A normative spatiotemporal MRI atlas of the fetal brain for automatic segmentation and analysis of early brain growth. *Sci. Rep.* **7**, 1–13. <https://doi.org/10.1038/s41598-017-00525-w> (2017).
- Gholipour, A. *et al.* Fetal MRI: A technical update with educational aspirations. *Concepts Magn. Reson. A Bridg. Educ. Res.* **43**, 237–266. <https://doi.org/10.1002/cmra.21321> (2014).
- Gholipour, A., Estroff, J. A. & Warfield, S. K. Robust super-resolution volume reconstruction from slice acquisitions: Application to fetal brain MRI. *IEEE Trans. Med. Imaging* **29**, 1739–1758. <https://doi.org/10.1109/TMI.2010.2051680> (2010).
- Rousseau, F., Kim, K., Studholme, C., Koob, M. & Dietemann, J. L. On super-resolution for fetal brain MRI. *Int. Conf. Med. Image Comput. Comput.-Assist. Interv.* **13**, 355–362. https://doi.org/10.1007/978-3-642-15745-5_44 (2010).
- Kuklisova-Murgasova, M., Quaghebeur, G., Rutherford, M. A., Hajnal, J. V. & Schnabel, J. A. Reconstruction of fetal brain MRI with intensity matching and complete outlier removal. *Med. Image Anal.* **16**, 1550–1564. <https://doi.org/10.1016/j.media.2012.07.004> (2012).
- Kainz, B. *et al.* Fast volume reconstruction from motion corrupted stacks of 2D slices. *IEEE Trans. Med. Imaging* **34**, 1901–1913. <https://doi.org/10.1109/TMI.2015.2415453> (2015).

16. Tourbier, S. *et al.* An efficient total variation algorithm for super-resolution in fetal brain MRI with adaptive regularization. *NeuroImage* **118**, 584. <https://doi.org/10.1016/j.neuroimage.2015.06.018> (2015).
17. Ebner, M. *et al.* An automated framework for localization, segmentation and super-resolution reconstruction of fetal brain MRI. *NeuroImage* **206**, 116324. <https://doi.org/10.1016/j.neuroimage.2019.116324> (2020).
18. Velasco-Annis, C. *et al.* Normative biometrics for fetal ocular growth using volumetric MRI reconstruction. *Prenat. Diagn.* **35**, 400–408. <https://doi.org/10.1002/pd.4558> (2015).
19. Pier, D. B. *et al.* 3D super-resolution motion-corrected MRI: Validation of fetal posterior fossa measurements. *J. Neuroimaging* **26**, 539–544. <https://doi.org/10.1111/jon.12342> (2016).
20. Khawam, M. *et al.* Fetal brain biometric measurements on 3D super-resolution reconstructed T2-weighted MRI: An intra- and inter-observer agreement study. *Front. Pediatr.* **9**, 651. <https://doi.org/10.3389/fped.2021.639746> (2021).
21. Makropoulos, A., Counsell, S. J. & Rueckert, D. A review on automatic fetal and neonatal brain MRI segmentation. *NeuroImage* **170**, 231–248. <https://doi.org/10.1016/j.neuroimage.2017.06.074> (2018).
22. Khalili, N. *et al.* Automatic brain tissue segmentation in fetal MRI using convolutional neural networks. *Magn. Reson. Imaging* **64**, 77–89. <https://doi.org/10.1016/j.mri.2019.05.020> (2019).
23. Hong, J. *et al.* Fetal cortical plate segmentation using fully convolutional networks with multiple plane aggregation. *Front. Neurosci.* **14**, 591683. <https://doi.org/10.3389/fnins.2020.591683> (2020).
24. Dou, H. *et al.* A deep attentive convolutional neural network for automatic cortical plate segmentation in fetal MRI. *IEEE Trans. Med. Imaging* **40**, 1123–1133. <https://doi.org/10.1109/TMI.2020.3046579> (2020).
25. Payette, K. *et al.* An automatic multi-tissue human fetal brain segmentation benchmark using the fetal tissue annotation dataset. *Sci. Data* **8**, 167. <https://doi.org/10.1038/s41597-021-00946-3> (2021).
26. Delannoy, Q. *et al.* SegSRGAN: Super-resolution and segmentation using generative adversarial networks—Application to neonatal brain MRI. *Comput. Biol. Med.* **120**, 103755. <https://doi.org/10.1016/j.compbiomed.2020.103755> (2020).
27. Wissmann, L., Santelli, C., Segars, W. P. & Kozerke, S. MRXCAT: Realistic numerical phantoms for cardiovascular magnetic resonance. *J. Cardiovasc. Magn. Reson.* **16**, 63. <https://doi.org/10.1186/s12968-014-0063-3> (2014).
28. Roy, C. W., Marini, D., Segars, W. P., Seed, M. & Macgowan, C. K. Fetal XCMR: A numerical phantom for fetal cardiovascular magnetic resonance imaging. *J. Cardiovasc. Magn. Reson.* **21**, 29. <https://doi.org/10.1186/s12968-019-0539-2> (2019).
29. Petersson, J. S., Christoffersson, J. O. & Golman, K. MRI simulation using the k-space formalism. *Magn. Reson. Imaging* **11**, 557–568. [https://doi.org/10.1016/0730-725x\(93\)90475-s](https://doi.org/10.1016/0730-725x(93)90475-s) (1993).
30. Xanthis, C. G., Venetis, I. E. & Aletras, A. H. High performance MRI simulations of motion on multi-GPU systems. *J. Cardiovasc. Magn. Reson.* **16**, 48. <https://doi.org/10.1186/1532-429X-16-48> (2014).
31. Liu, F., Velikina, J. V., Block, W. F., Kijowski, R. & Samsonov, A. A. Fast realistic MRI simulations based on generalized multi-pool exchange tissue model. *IEEE Trans. Med. Imaging* **36**, 527–537. <https://doi.org/10.1109/TMI.2016.2620961> (2017).
32. Drobniak, I., Gavaghan, D., Süli, E., Pitt-Francis, J. & Jenkinson, M. Development of a functional magnetic resonance imaging simulator for modeling realistic rigid-body motion artifacts. *Magn. Reson. Med.* **56**, 364–380. <https://doi.org/10.1002/mrm.20939> (2006).
33. Wang, Z., Bovik, A., Sheikh, H. & Simoncelli, E. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612. <https://doi.org/10.1109/TIP.2003.819861> (2004).
34. Sara, U., Akter, M. & Uddin, M. S. Image quality assessment through FSIM, SSIM, MSE and PSNR—A comparative study. *J. Comput. Commun.* **7**, 8–18. <https://doi.org/10.4236/jcc.2019.73002> (2019).
35. Shepp, L. & Logan, B. The Fourier reconstruction of a head section. *IEEE Trans. Nucl. Sci.* **21**, 21–43. <https://doi.org/10.1109/TNS.1974.6499235> (1974).
36. Stöcker, T., Vahedipour, K., Pflugfelder, D. & Shah, N. J. High-performance computing MRI simulations. *Magn. Reson. Med.* **64**, 186–193. <https://doi.org/10.1002/mrm.22406> (2010).
37. Xanthis, C. G., Venetis, I. E., Chalkias, A. V. & Aletras, A. H. MRISIMUL: A GPU-based parallel approach to MRI simulations. *IEEE Trans. Med. Imaging* **33**, 607–617. <https://doi.org/10.1109/TMI.2013.2292119> (2014).
38. Hennig, J., Weigel, M. & Scheffler, K. Calculation of flip angles for echo trains with predefined amplitudes with the extended phase graph (EPG)-algorithm: Principles and applications to hyperecho and TRAPS sequences. *Magn. Reson. Med.* **51**, 68–80. <https://doi.org/10.1002/mrm.10658> (2004).
39. Weigel, M. Extended phase graphs: Dephasing, RF pulses, and echoes—Pure and simple. *J. Magn. Reson. Imaging* **41**, 266–295. <https://doi.org/10.1002/jmri.24619> (2015).
40. Malik, S. J., Teixeira, R. P. A. G. & Hajnal, J. V. Extended phase graph formalism for systems with magnetization transfer and exchange. *Magn. Reson. Med.* **80**, 767–779. <https://doi.org/10.1002/mrm.27040> (2018).
41. Hennig, J. Multiecho imaging sequences with low refocusing flip angles. *J. Magn. Reson.* **1969**(78), 397–407. [https://doi.org/10.1016/0022-2364\(88\)90128-X](https://doi.org/10.1016/0022-2364(88)90128-X) (1988).
42. Hennig, J. & Scheffler, K. Hyperechoes. *Magn. Reson. Med.* **46**, 6–12. <https://doi.org/10.1002/mrm.1153> (2001).
43. Hennig, J., Weigel, M. & Scheffler, K. Multiecho sequences with variable refocusing flip angles: Optimization of signal behavior using smooth transitions between pseudo steady states (TRAPS). *Magn. Reson. Med.* **49**, 527–535. <https://doi.org/10.1002/mrm.10391> (2003).
44. Weigel, M. & Hennig, J. Contrast behavior and relaxation effects of conventional and hyperecho-turbo spin echo sequences at 1.5 and 3 T. *Magn. Reson. Med.* **55**, 826–835. <https://doi.org/10.1002/mrm.20816> (2006).
45. Weigel, M., Schwenk, S., Kiselev, V. G., Scheffler, K. & Hennig, J. Extended phase graphs with anisotropic diffusion. *J. Magn. Reson.* **205**, 276–285. <https://doi.org/10.1016/j.jmr.2010.05.011> (2010).
46. Tamir, J. I. *et al.* Computational MRI with physics-based constraints: Application to multicontrast and quantitative imaging. *IEEE Signal Process. Mag.* **37**, 94–104. <https://doi.org/10.1109/MSP.2019.2940062> (2020).
47. Sbrizzi, A. *et al.* Optimal control design of turbo spin-echo sequences with applications to parallel-transmit systems. *Magn. Reson. Med.* **77**, 361. <https://doi.org/10.1002/mrm.26084> (2016).
48. Keerthivasan, M. B. *et al.* An efficient 3D stack-of-stars turbo spin echo pulse sequence for simultaneous T2-weighted imaging and T2 mapping. *Magn. Reson. Med.* **82**, 326–341. <https://doi.org/10.1002/mrm.27737> (2019).
49. Prasloski, T., Mädler, B., Xiang, Q.-S., MacKay, A. & Jones, C. Applications of stimulated echo correction to multicomponent T2 analysis. *Magn. Reson. Med.* **67**, 1803–1814. <https://doi.org/10.1002/mrm.23157> (2012).
50. Lankford, C. L., Dortch, R. D. & Does, M. D. Fast T2 mapping with multiple echo, Caesar cipher acquisition and model-based reconstruction. *Magn. Reson. Med.* **73**, 1065–1074. <https://doi.org/10.1002/mrm.25221> (2015).
51. Cloos, M. A. *et al.* Multiparametric imaging with heterogeneous radiofrequency fields. *Nat. Commun.* **7**, 12445. <https://doi.org/10.1038/ncomms12445> (2016).
52. Tourbier, S., De Dumast, P., Kebiri, H., Hagmann, P. & Bach Cuadra, M. Medical-Image-Analysis-Laboratory/mialsuperresolutiontoolkit: MIAL Super-Resolution Toolkit v2.0.1. *Zenodo*. <https://doi.org/10.5281/zenodo.4392788> (2020).
53. Hagmann, C. F. *et al.* T2 at MR imaging is an objective quantitative measure of cerebral white matter signal intensity abnormality in preterm infants at term-equivalent age. *Radiology* **252**, 209–217. <https://doi.org/10.1148/radiol.2522080589> (2009).
54. Nossin-Manor, R. *et al.* Quantitative MRI in the very preterm brain: Assessing tissue organization and myelination using magnetization transfer, diffusion tensor and T1 imaging. *NeuroImage* **64**, 505–516. <https://doi.org/10.1016/j.neuroimage.2012.08.086> (2013).

55. Blazejewska, A. I. *et al.* 3D in utero quantification of T2* relaxation times in human fetal brain tissues for age optimized structural and functional MRI. *Magn. Reson. Med.* **78**, 909–916. <https://doi.org/10.1002/mrm.26471> (2017).
56. Yarnykh, V. L., Prihod'ko, I. Y., Savelov, A. A. & Korostyshevskaya, A. M. Quantitative assessment of normal fetal brain myelination using fast macromolecular proton fraction mapping. *Am. J. Neuroradiol.* **39**, 1341–1348. <https://doi.org/10.3174/ajnr.A5668> (2018).
57. Vasylechko, S. *et al.* T2* relaxometry of fetal brain at 1.5 Tesla using a motion tolerant method. *Magn. Reson. Med.* **73**, 1795–1802. <https://doi.org/10.1002/mrm.25299> (2015).
58. Stanisz, G. J. *et al.* T1, T2 relaxation and magnetization transfer in tissue at 3T. *Magn. Reson. Med.* **54**, 507–512. <https://doi.org/10.1002/mrm.20605> (2005).
59. Rooney, W. D. *et al.* Magnetic field and tissue dependencies of human brain longitudinal ¹H₂O relaxation in vivo. *Magn. Reson. Med.* **57**, 308–318. <https://doi.org/10.1002/mrm.21122> (2007).
60. Shin, W., Gu, H. & Yang, Y. Fast high-resolution T1 mapping using inversion-recovery look-locker echo-planar imaging at steady state: Optimization for accuracy and reliability. *Magn. Reson. Med.* **61**, 899–906. <https://doi.org/10.1002/mrm.21836> (2009).
61. Bojorquez, J. Z. *et al.* What are normal relaxation times of tissues at 3 T? *Magn. Reson. Imaging* **35**, 69–80. <https://doi.org/10.1016/j.mri.2016.08.021> (2017).
62. Daoust, A. *et al.* Transverse relaxation of cerebrospinal fluid depends on glucose concentration. *Magn. Reson. Imaging* **44**, 72–81. <https://doi.org/10.1016/j.mri.2017.08.001> (2017).
63. Cocosco, C. A., Kollokian, V., Kwan, R.K.-S., Pike, G. B. & Evans, A. C. BrainWeb: Online interface to a 3D MRI simulated brain database. *Neuroimage* **5**, 425 (1997).
64. BrainWeb: Simulated Brain Database. <https://brainweb.bic.mni.mcgill.ca/brainweb/>. (Accessed 30 Sept 2020).
65. Kwan, R., Evans, A. & Pike, B. An extensible MRI simulator for post-processing evaluation. *Vis. Biomed. Comput.* <https://doi.org/10.1007/BFb0046947> (1996).
66. Collins, D. L. *et al.* Design and construction of a realistic digital brain phantom. *IEEE Trans. Med. Imaging* **17**, 463–468. <https://doi.org/10.1109/42.712135> (1998).
67. Kwan, R. K.-S., Evans, A. C. & Pike, G. B. MRI simulation-based evaluation of image-processing and classification methods. *IEEE Trans. Med. Imaging* **18**, 1085–1097. <https://doi.org/10.1109/42.816072> (1999).
68. Busse, R. F., Hariharan, H., Vu, A. & Brittain, J. H. Fast spin echo sequences with very long echo trains: Design of variable refocusing flip angle schedules and generation of clinical t2 contrast. *Magn. Reson. Med.* **55**, 1030–1037. <https://doi.org/10.1002/mrm.20863> (2006).
69. Lowe, M. J. & Sorenson, J. A. Spatially filtering functional magnetic resonance imaging data. *Magn. Reson. Med.* **37**, 723–729. <https://doi.org/10.1002/mrm.1910370514> (1997).
70. Friedman, L., Glover, G. H., Krenz, D. & Magnotta, V. Reducing inter-scanner variability of activation in a multicenter fMRI study: Role of smoothness equalization. *Neuroimage* **32**, 1656–1668. <https://doi.org/10.1016/j.neuroimage.2006.03.062> (2006).
71. Rousseau, F. *et al.* Registration-based approach for reconstruction of high-resolution in utero fetal MR brain images. *Acad. Radiol.* **13**, 1072–1081. <https://doi.org/10.1016/j.acra.2006.05.003> (2006).
72. Oubel, E., Koob, M., Studholme, C., Dietemann, J.-L. & Rousseau, F. Reconstruction of scattered data in fetal diffusion MRI. *Med. Image Anal.* **16**, 28–37. <https://doi.org/10.1016/j.media.2011.04.004> (2012).
73. Lajous, H., Ledoux, J.-B., Hilbert, T., van Heeswijk, R. B. & Bach Cuadra, M. Dataset T2 mapping from super-resolution-reconstructed clinical fast spin echo magnetic resonance acquisitions. *Zenodo*. <https://doi.org/10.5281/zenodo.3931812> (2020).
74. Yushkevich, P. A. *et al.* User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* **31**, 1116–1128. <https://doi.org/10.1016/j.neuroimage.2006.01.015> (2006).
75. Johnson, H. J., Harris, G. & Williams, K. BRAINSFit: Mutual information registrations of whole-brain 3D images, using the insight toolkit. *Insight J.* **57**, 1–10 (2007).
76. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Lecture Notes in Computer Science* (eds Navab, N. *et al.*) 234–241 (Springer, 2015).
77. Milletari, F., Navab, N. & Ahmadi, S.-A. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 Fourth International Conference on 3D Vision (3DV)*, 565–571, <https://doi.org/10.1109/3DV.2016.79> (2016).
78. Abadi, M. *et al.* TensorFlow: A system for large-scale machine learning. In *Proc. 12th USENIX Conference on Operating Systems Design and Implementation*, OSDI'16, 265–283 (USENIX Association, 2016).
79. Pérez-García, F., Sparks, R. & Ourselin, S. Torchio: A python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Comput. Methods Progr. Biomed.* <https://doi.org/10.1016/j.cmpb.2021.106236> (2021).
80. Dice, L. R. Measures of the amount of ecologic association between species. *Ecology* **26**, 297–302. <https://doi.org/10.2307/1932409> (1945).
81. Lajous, H. *et al.* A magnetic resonance imaging simulation framework of the developing fetal brain. In *Proc. 29th Annual Meeting of ISMRM* (virtual, 2021). Program number 0734.
82. Christ, A. *et al.* The Virtual Family—Development of surface-based anatomical models of two adults and two children for dosimetric simulations. *Phys. Med. Biol.* **55**, N23–N38. <https://doi.org/10.1088/0031-9155/55/2/N01> (2010).
83. Gosselin, M.-C. *et al.* Development of a new generation of high-resolution anatomical models for medical device evaluation: The virtual population 3.0. *Phys. Med. Biol.* **59**, 5287–5303. <https://doi.org/10.1088/0031-9155/59/18/5287> (2014).
84. Hasgall, P. A. *et al.* IT'IS Database for thermal and electromagnetic parameters of biological tissues, Version 4.0. <https://doi.org/10.13099/VIP21000-04-0> (2018).
85. Segars, W. & Tsui, B. MCAT to XCAT: The evolution of 4-D computerized phantoms for imaging research. *Proc. IEEE* **97**, 1954–1968. <https://doi.org/10.1109/JPROC.2009.2022417> (2010).
86. Segars, W. P., Sturgeon, G., Mendonca, S., Grimes, J. & Tsui, B. M. W. 4D XCAT phantom for multimodality imaging research. *Med. Phys.* **37**, 4902–4915. <https://doi.org/10.1118/1.3480985> (2010).
87. Norris, H. *et al.* A set of 4D pediatric XCAT reference phantoms for multimodality research. *Med. Phys.* **41**, 033701. <https://doi.org/10.1118/1.4864238> (2014).
88. Lajous, H. *et al.* Dataset A Fetal Brain magnetic resonance Acquisition Numerical phantom (FaBIAN). *Zenodo*. <https://doi.org/10.5281/zenodo.6477946> (2022).
89. Lajous, H., Roy, C. W., Yerly, J. & Bach Cuadra, M. FaBIAN v1.2. *Zenodo*. <https://doi.org/10.5281/zenodo.5471094> (2022).

Acknowledgements

This work was supported by the Swiss National Science Foundation through Grants 182602, 141283 and 173129, by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie project TRABIT (Agreement No 765148), by the Forschungszentrum für das Kind (FZK), the Hasler Foundation, the Anna Muller–Grocholski Foundation, and the Zurich Neuroscience Center PhD Grant. We acknowledge access to the facilities and expertise of the CIBM Center for Biomedical Imaging, a Swiss research center of

excellence founded and supported by Lausanne University Hospital (CHUV), University of Lausanne (UNIL), Ecole polytechnique fédérale de Lausanne (EPFL), University of Geneva (UNIGE) and Geneva University Hospitals (HUG). We would like to thank Dr. Ruth Tuura from the Center for MR Research of University Children's Hospital Zurich (Zurich, Switzerland) for her support in answering questions about acquisition parameters specific to the implementation of the SS-FSE sequence by GE. We would also like to thank Ji Hui from the Center for MR Research of University Children's Hospital Zurich (Zurich, Switzerland) for the annotation of the SR fetal brain MR images in subjects scanned at Kispi.

Author contributions

H.L. investigated the HASTE sequence pulse design using the IDEA sequence programming interface provided by Siemens Healthineers, built on the initial simulation framework to develop FaBiAN, implemented additional features such as the inclusion of intensity non-uniformity fields and stochastic motion, and extended this numerical phantom, originally thought to simulate HASTE acquisitions of the fetal brain, to FSE sequences from other MR vendors. H.L. suggested a classification of segmented brain tissues as gray matter, white matter or cerebrospinal fluid, reviewed the state of the art in order to find out T1, T2 and T2* values reported in the developing fetal brain as well as in premature and full-term newborns at 1.5 T, and extrapolated from the more general literature the corresponding T1 and T2 values at 3 T. H.L. automatically generated the brain masks of the simulated images and manually refined them, reconstructed the super-resolution fetal brain volumes from clinical, respectively synthetic, 2D motion-corrupted series, assessed the motion level in all clinical series, participated in the design of the different case studies, analyzed all the results, prepared the tables and figures, and wrote the article. C.W.R. implemented the skeleton of the original simulator, helped H.L. in optimizing it, and suggested complementary case studies. T.H. helped H.L. to understand the design of the HASTE sequence and to improve the simulation process. C.W.R. and T.H. equally contributed to this work. P.d.D. provided clinical cases acquired at CHUV from the fetal MRI database whose SR reconstructions were previously evaluated as excellent. P.d.D. and H.K. participated in the design of the data augmentation experiments using simulated images in the context of fetal brain tissue segmentation. P.d.D. implemented the deep learning network for fetal brain tissue segmentation and evaluated its performance in the different configurations studied. P.d.D. and S.T. assisted H.L. in using the SR reconstruction pipeline⁵². S.T. especially helped with the installation procedure. Y.A.G. provided a script that saves matlab matrices to nifti images with the right orientation matrix. J.Y. provided a script that preprocesses the data in k-space using a Fermi low-pass filter before zero-interpolation filling. T.Y. provided a script that computes a motion index from the brain masks. K.P. shared the data from fifteen subjects scanned at Kispi, including clinical LR SS-FSE images of the fetal brain and corresponding SR reconstructions together with their manual annotations. K.P. performed the SR reconstructions of the data from Kispi. K.P. also provided information about the protocol for fetal brain examination as it is conducted in clinical routine at Kispi. J.B.L. provided information on the clinical protocol for fetal brain MRI as it is performed at CHUV. R.M. assessed the clinical cases. A.J. supported the use of detailed clinical cases from Kispi to extend our simulation framework to FSE sequences from other MR vendors and demonstrate its added value in providing multiple synthetic data at no cost for data augmentation strategies. M.K. and V.D. validated and complemented the labels of the various anatomical structures in the developing fetal brain. V.D. provided approximate relaxometric T1 and T2 values in the cerebrospinal fluid in pediatric patients. M.K. detailed the features radiologists routinely analyze in clinical images of the fetal brain to determine the neurotypical or pathological status of the fetus, as well as the various artefacts they typically encounter in LR HASTE images. M.K. and V.D. provided qualitative assessment of the simulated HASTE images. P.H., T.K. and M.S. supported the writing of this manuscript and provided regular feedback during the project. M.B.C. regularly assessed the realistic appearance of the simulated HASTE images in comparison with typical clinical acquisitions of the fetal brain at CHUV throughout the implementation of the simulation pipeline, qualitatively estimated the level of motion in clinical series, suggested some application examples where the simulator could be particularly valuable, participated in the design of the various experiments presented in this manuscript, and supported the analysis of the results as well as the writing of the manuscript. M.B.C. is the recipient of the FNS grant that supports the whole project. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-10335-4>.

Correspondence and requests for materials should be addressed to H.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022