Biomimetic phantom with anatomical accuracy for evaluating brain volumetric measurements with magnetic resonance imaging

Mehran Azimbagirad[a,b,*,†] Felipe Wilker Grillo[b,†] Yaser Hadadian[b] Antonio Adilton Oliveira Carneiro[b] and Luiz Otavio Murta Jr.[c]

[a]University of Western Brittany, Faculty of Medicine and Health Sciences, Brest, France
[b]University of São Paulo, Department of Physics, Faculty of Philosophy, Science and Languages, Ribeirão Preto, São Paulo, Brazil
[c]University of São Paulo, Department of Computing and Mathematics, Faculty of Philosophy, Science and Languages, Ribeirão Preto, São Paulo, Brazil

Abstract

Purpose: Brain image volumetric measurements (BVM) methods have been used to quantify brain tissue volumes using magnetic resonance imaging (MRI) when investigating abnormalities. Although BVM methods are widely used, they need to be evaluated to quantify their reliability. Currently, the gold-standard reference to evaluate a BVM is usually manual labeling measurement. Manual volume labeling is a time-consuming and expensive task, but the confidence level ascribed to this method is not absolute. We describe and evaluate a biomimetic brain phantom as an alternative for the manual validation of BVM.

Methods: We printed a three-dimensional (3D) brain mold using an MRI of a three-year-old boy diagnosed with Sturge-Weber syndrome. Then we prepared three different mixtures of styrene-ethylene/butylene-styrene gel and paraffin to mimic white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). The mold was filled by these three mixtures with known volumes. We scanned the brain phantom using two MRI scanners, 1.5 and 3.0 Tesla. Our suggestion is a new challenging model to evaluate the BVM which includes the measured volumes of the phantom compartments and its MRI. We investigated the performance of an automatic BVM, i.e., the expectation–maximization (EM) method, to estimate its accuracy in BVM.

Results: The automatic BVM results using the EM method showed a relative error (regarding the phantom volume) of 0.08, 0.03, and 0.13 ($\pm$0.03 uncertainty) percentages of the GM, CSF, and WM volume, respectively, which was in good agreement with the results reported using manual segmentation.

Conclusions: The phantom can be a potential quantifier for a wide range of segmentation methods.

1 Introduction

Medical imaging techniques seek to reveal organs' internal structure to diagnose disorders and follow up their behavior in response to treatments. These techniques include two central steps, i.e., hardware setup and image processing algorithms. Alongside prior information and expertise,

---

*Address all correspondence to Mehran Azimbagirad, mehran.azimbagirad@univ-brest.fr

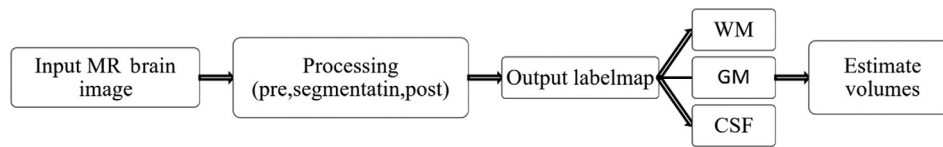†These authors contributed equally to this work

Fig. 1 A pipeline for volume estimation using the image segmentation method. In this schematic, the volumes of three main brain compartments (WM, GM, and CSF) are estimated using a label map.

radiologists use image processing packages to interpret the extracted results to suggest a patient's treatment strategy under examination. Therefore, the precision of image processing techniques is of crucial importance for the treatment.

Brain image segmentation is one of the image processing techniques used for visualization[1] and volumetric analysis[2–4] of the brain compartments as well as the possibility to detect lesions[5,6] or tumors.[7,8] Figure 1 presents a common pipeline for volumetric brain segmentation. The pipeline produces a label map that classifies each image voxel into tissue categories, i.e., white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). Thus, one can calculate the volume of compartments using the labels and the image's voxel size.

Several volumetric brain segmentation methods have been proposed using different techniques, e.g., expectation–maximization (EM) algorithm,[9] K-mean,[10] and q-entropy[11] for either estimating a specific region or the total brain volume. Although the number of such techniques has proliferated as research groups seek new methods with improved accuracy and faster execution, one needs to evaluate their performance to quantify the reliability of the extracted results from those methods.

It is common to use similarity coefficients[12] to evaluate the label map's precision produced by an image segmentation method. In this type of evaluation, a reference image (gold standard) is needed to calculate the similarity coefficient. Therefore, an index is calculated to show how much the label map is similar to the gold standard. If the index approaches 1, a higher similarity is achieved, and if the index is closer to 0, a lower similarity is concluded. The evaluation pipeline using the similarity metric is illustrated in Fig. 2.

To date, there are two sources of gold standard references to calculate the similarity metrics: expert manual segmentation and phantom imaging. Manual segmentation is the most reliable one, but it is time-consuming and strongly operator dependent.[13] One can use postmortem volume quantification as a reference to access the accuracy of segmentation methods, though tissue deformation prevents voxel-wise comparison.[14] There are significant visual ambiguities that make the manual segmentation a very challenging task. For instance, the contrast range of the human eyes is limited and may not distinguish voxels with very similar intensities, as shown in Fig. 3. As an alternative for this infrequent source, phantoms are more naturally accessible sources.
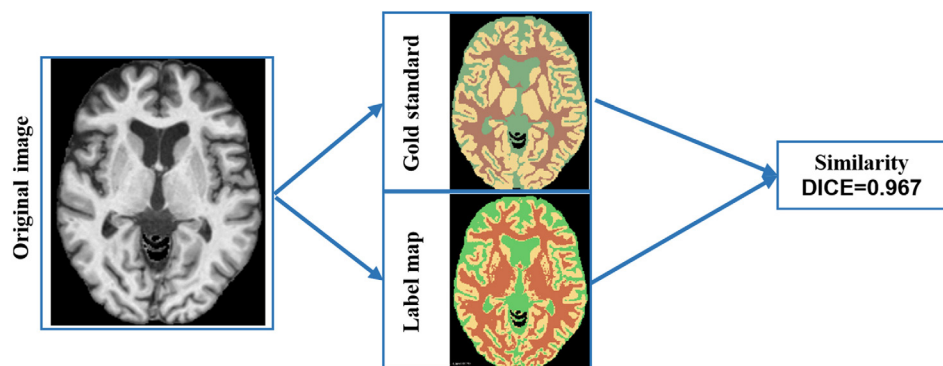

Fig. 2 Evaluation of a brain segmentation method using DICE (a similarity coefficient). The calculated coefficient shows the similarity between label maps and the gold standard.
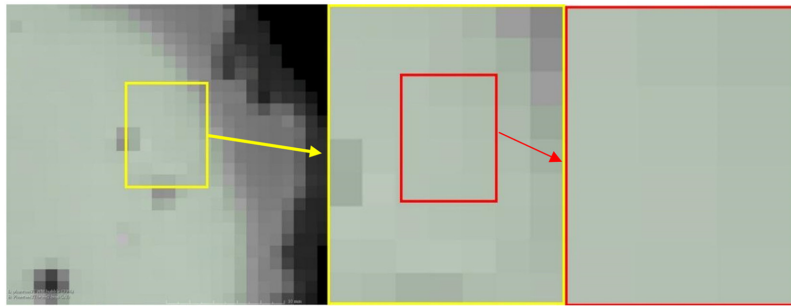
Fig. 3 A part of the axial view of a brain image. Yellow rectangular shows a part of the MRI including two tissues. Red rectangular shows 12 pixels with different intensities. It is difficult for human eyes to choose a very similar pixel in the border for labeling because eyes may not be able to distinguish the difference between them.

Phantoms are divided into digital[15] and physical[16] categories. Digital phantoms have been widely used for evaluating brain segmentation methods more than the others as ground truth.[17–19] The first and most important advantage of digital phantoms is the controllability of the produced image features. This means that by changing the features of the simulation[20] in digital phantoms, the drawbacks of a volumetric segmentation algorithm can be revealed. Reproducing this kind of phantoms is cost-effective and not a time-consuming task. Magnetic resonance imaging (MRI) brain simulations usually mimic healthy brains [Fig. 4(a)] and do not contain all the artifacts that may exist in real situations [Fig. 4(c)]. Artifacts such as noise and magnetic field bias effects are added to digital phantoms to produce a plausible simulation close to an actual brain image using different algorithms. However, one can easily remove the artifacts by recognizing their type.

A physical phantom, also known as a clinical simulator, is an object made by a tissue-mimicking material (TMM) with tunable properties. TMMs can simulate different properties in different imaging modalities to evaluate, analyze, and tune various imaging devices.[22] Physical phantoms can mimic most human body parts, either addressing surgical planning such as cardiac,[23] oral surgery,[24] and breast[25,26] or simulating needle-based interventions.[27] Recently, phantoms with desired imaging characteristics for three-dimensional (3D) techniques such as computer tomography (CT) and MRI[28] and ultrasound[29,30] have been employed.

Although the physical phantoms mimicking the brain are not new,[31,32] recently, they have been used and developed for brain image analysis.[33–38] They can also be used to investigate and characterize imaging devices, e.g., photoacoustic imaging systems[39] and diffuse imaging and spectroscopy.[40] Here, the acquired images will be analyzed to adjust the imaging techniques' parameters to achieve the best possible image quality. Construction time and cost are disadvantages of physical phantoms compared to digital phantoms. A significant difference between
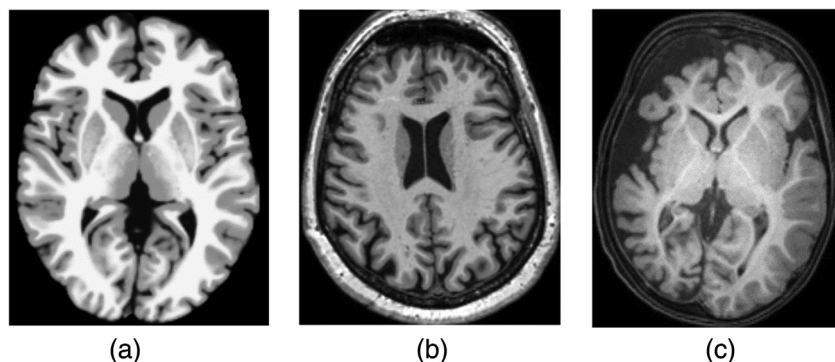

Fig. 4 (a) Simulated brain;[17] (b) healthy brain; and (c) patient brain with Sturge-Weber disease.[21] Atrophy can be seen on the left side of the patient's brain.

physical and digital phantoms is the presence of real artifacts in physical ones. Despite all the pros and cons of physical phantoms, many companies produce them for various research fields and educational purposes. For brevity, we will use the term phantom to indicate a physical phantom throughout the rest of this paper.

A wide range of materials can be used to construct a phantom. Two of the most used classes in TMMs are natural polymers (e.g., gelatin and agar) and synthetic polymers [e.g., polyvinyl alcohol (PVA) cryogel and polyacrylamide[41]]. Cabrelli et al.[42] provided a mixture of the copolymer in oil styrene-ethylene/butylene-styrene (SEBS), i.e., kraton polymers/paraffin wax, in order to mimic the main brain compartments. This mixture in MRI has a specific range of intensity in which, by changing the ratio of paraffin wax to SEBS gel, the intensities can be adjusted similar to the intensities of the brain compartments in MRI. In addition, the phantoms made of such materials have significantly lower temporal degradation compared to water-based phantoms.[43,44]

Three-dimensional models can be produced based on CT or MRI volumetric medical images.[28] This 3D printed model is used as a mold in the phantom preparation process.[24] In this work, we prepared a 3D brain model and used it as the mold (consisting of two separated molds for left and right lobes) filled with three mixtures of SEBS gel and paraffin to mimic WM, GM, and CSF in the brain. Since each phantom compartment's volume before filling the molds was measured experimentally using two different methods, it can be compared with the measured volumes by any brain image volumetric measurements (BVM). Therefore, this phantom and the measures of its compartments are our suggestions as a quantifier for BVMs. As an implementation example, our proposed model was used to evaluate EM[45] to assess its accuracy when estimating the phantom brain compartments.

2 Materials and Methods

2.1 Material Preparation

We used nine mixtures of the copolymer (SEBS) in mineral oil in a concentration of 10% w/w and paraffin wax to choose three among them mimicking the brain tissues. Paraffin wax was added in a mass fraction of 0% to 80% (by 10% increments) of solvent (oil mass) to the SEBS gel.[42] The sample preparation of the SEBS copolymer, mineral oil, and paraffin was continually stirred. The mixtures remained at 120°C for ∼5 h in an oven (YK Oven vac 64L) to ensure complete homogenization.

2.2 Phantom Preparation

A pilot phantom was constructed using a sample consisting of a mixture of SEBS gel/paraffin (90% w/w) and a hypothalamus morphology mold from a didactical model.[46] We scanned the hypothalamus phantom in a 3.0 Tesla MRI scanner to reveal the inhomogeneity and intensity range. Next, we covered the pilot phantom with a sample layer with a concentration of 70% and then one more layer of 90% w/w SEBS gel/paraffin. The covered hypothalamus phantom with the nine cylindrical samples ($20 \times 20 \times 20$ mm^3) were scanned by 1.5 and 3.0 Tesla scanners into three samples for manufacturing the brain compartments. A brain MRI of a 3-year-old boy diagnosed with Sturge-Weber syndrome[21] was used to print a 3D silicone mold (as a two-piece model for the lobes of the brain). This disease results in an asymmetric lobe, unlike common models, which are symmetric and homogenous. These geometric properties provide a challenging model for evaluating BVM. The molds were printed (polylactic acid fused filament) separately, using a Zmorph 2.0 S (LLC, Wroclaw, Poland) printer with a z-layer: 0.1 mm, path width 0.4 mm. Then both molds were filled using the first sample [100% SEBS gel, 450.320 g (±0.005) or 531,300 mm^3 (±0.01%)] as WM. After extracting the WM mimicking phantoms (two lobes) from the molds, they were covered using the second sample [70% w/w SEBS gel/paraffin, 125.410 g (±0.005) or 147,100 mm^3 (±0.01%)] as GM. Afterward, these two WM–GM phantoms were joined together and covered with the third sample [20% w/w SEBS gel/paraffin, 171.270 g (±0.005) or 199,100 mm^3 (±0.01%)] as CSF.

2.3 Error Analysis

To determine the exact volume of each part of the phantom, we estimated the samples' density by two different methods. Using the well-known water displacement method and using an aluminum cylindrical mold with known volume (5 ± 0.001 mm inner radius and 10 ± 0.001 mm height), which was filled with the sample. In both methods, an analytical balance with a readability of 0.000001 g was employed. However, to measure the mass of the materials to fill the molds, we used a balance with ± 0.005 g uncertainty. There are two types of uncertainties to measure the volumes: instrument error (systematic) and formula's error (computational). For the instruments, i.e., balances and digital micrometer, the systematic errors were mentioned. For the density formula, we used the fractional uncertainties as

$$\rho = \frac{m}{V} \rightarrow \frac{\delta\rho}{\rho} = \sqrt{\left(\frac{\delta m}{m}\right)^2 + \left(\frac{\delta V}{V}\right)^2}, \tag{1}$$

but for the volume of the aluminum cylindrical mold

$$V = \pi \times r^2 \times h \rightarrow \frac{\delta V}{V} \cong \sqrt{2\left(\frac{\delta r}{r}\right)^2 + \left(\frac{\delta h}{h}\right)^2} = \sqrt{2\left(\frac{0.001}{5}\right)^2 + \left(\frac{0.001}{10}\right)^2}$$

$$\cong 0.0003 \text{ or } 0.03\%. \tag{2}$$

Therefore, for the density, we have

$$\frac{\delta\rho}{\rho} = \sqrt{\left(\frac{10^{-6}}{2}\right)^2 + (0.0003)^2} \cong 0.0003 \text{ or } 0.03\%, \tag{3}$$

and for the volume of the phantom compartments

$$V = \frac{m}{\rho} \rightarrow \frac{\delta V}{V} = \sqrt{\left(\frac{\delta m}{m}\right)^2 + \left(\frac{\delta\rho}{\rho}\right)^2}. \tag{4}$$

For example, for GM volume estimation in the lab, total error in percentage was calculated as follows:

$$\frac{\delta V_{GM}}{V_{GM}} = \sqrt{\left(\frac{5}{125410}\right)^2 + \left(\frac{\delta\rho}{\rho}\right)^2} \cong \sqrt{0.00004^2 + 0.0003^2} \cong 0.03\%. \tag{5}$$

We estimated the maximum error (systematic and computational) for WM or CSF using Eq. (5) as 0.03% of the total volume.

2.4 MRI Characterization

We scanned the material samples and the phantoms using two different MRI scanners, i.e., a 1.5-T scanner and a higher magnetic field strength 3.0-T scanner. The details of both scanners can be seen in Table 1.

2.5 BVM Evaluation

The acquired image and the phantom's volumes were used as a new challenging model for evaluating BVM. This model can be used to evaluate any BVM, e.g., a brain volume segmentation method by following these instructions.

Step 1. Use the brain phantom image as the input image.
Step 2. Use a BVM to segment the image into its compartments and extract a label map.
Step 3. Use the label map to calculate the volumes of the brain phantom compartments.

Table 1 MRI scanners specification used in this study.

Specifications	Scanner 1	Scanner 2
Manufacturer (model)	Philips Medical Systems (Achieva)	Philips Medical Systems (Achieva)
Magnetic field strength (Tesla)	1.5	3.0
Receiving coil	SENSE-Head-8	Dual coil
Transmitting coil	B	B
Scanning sequence	GR	GR
Sequence variant	MP	MP
Repetition time (ms)	8.04	7
Echo time (ms)	3.73	3.21
Flip angle (deg)	8	8
Acquisition matrix	$400 \times 400 \times 160$	$240 \times 240 \times 160$
Resolution (voxels per mm)	$0.6 \times 0.6 \times 1.0$	$1.0 \times 1.0 \times 1.0$
Protocol name	SAG_T1_3D SENSE	3DT1_VBM SENSE

Step 4. Calculate the error using the values calculated in step 3 and the volumes measured experimentally.

We evaluated the EM segmentation as a BVM using this instruction. The software to perform EM segmentation and calculate the volumes was a 3DSlicer.[47] Two relative errors were calculated for each phantom compartment, i.e., relative error with respect to phantom volume (RE1) and relative error with respect to each compartment (RE2).

3 Results

Figure 5 shows the MRI image of the hypothalamus phantom made by 90% w/w SEBS gel/paraffin wax, covered by 30% and 10% paraffin, respectively, in which we can see a few bubbles in the sagittal (yellow) view of the hypothalamus (down left). Nine mixtures of SEBS gel and paraffin wax are surrounding the hypothalamus.

Table 2 presents the range intensities of the samples in both scanners calculated by 3DSlicer software. The range intensities of samples in 1.5 and 3.0 Tesla devices are different indirectly due to time repetition (TR) and time echo (TE), but directly due to T1 relaxometry differences (which depend on the B0 field strength).

We compared the range intensity of three real MRIs with the samples to choose three samples among the mixtures for manufacturing the brain compartments. Table 3 presents the intensity statistics of the three real images.

Since different scanners were used to acquire the images, one cannot expect to obtain the same range of intensities, even with the same modality protocols. Nevertheless, by scaling the intensities of the samples and the real images from 0 to 1 we chose samples 1, 4, and 9, which had maximum similarity to CSF (~0 to 0.3), GM (~0.3 to 0.8), and WM (~0.8 to 1.0) in real images.

We used sample 1 (100% SEBS gel as WM) to fill the two-piece model lobes. Then sample 4 (70% SEBS gel with 30% paraffin as GM) was used to cover the lobes. In Fig. 6(a), a few bright (noises) and dark (bubbles) spots have appeared. The two lobes in the previous step were attached and covered with sample 9 (20% SEBS gel 80% paraffin wax as CSF) to complete the brain phantom with three compartments. In addition to bubbles and noises, the bias field effect can also be seen in this image. The bias field images produced by N4Bias Field Correction module in the 3DSlicer are shown in Fig. 6(b).
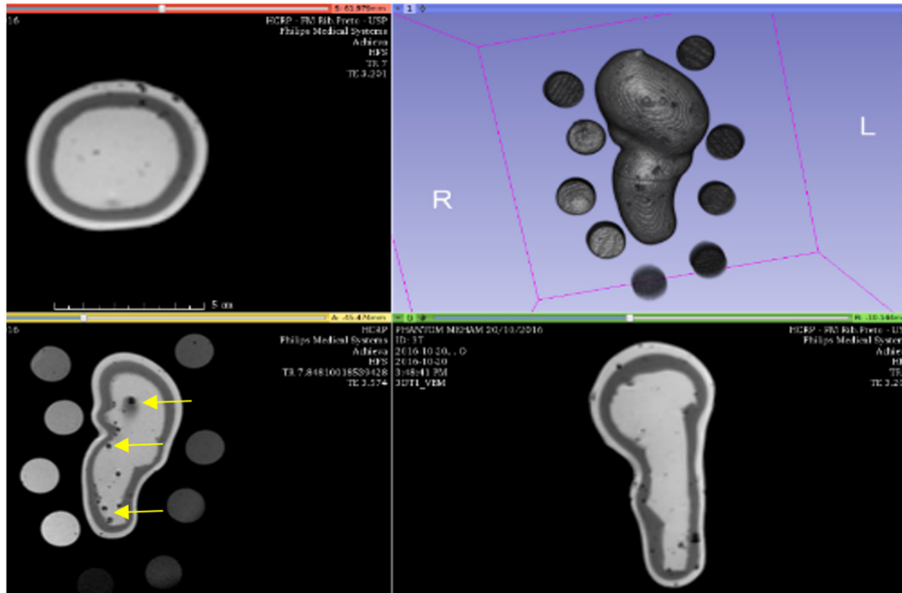
Fig. 5 Hypothalamus phantom made by one sample (90% SEBS gel mixed to 10% paraffin wax) covered by 30% and 10% paraffin, respectively. Top left axial, right 3D, down left the sagittal and downright coronal view of the acquired image in 3DSlicer software. Nine different combinations of SEBS gel and paraffin wax are surrounding a hypothalamus. The yellow arrows show the bubbles.

After manufacturing the entire brain phantom, the total volume was measured (here only using water displacement, because using density and mass was not applicable) again to determine the difference between before and after attaching the three samples. The total volume of the final product was more than the sum of all volumes separately. Indeed, this difference was due to the presence of the undesired bubbles. The final volume of the phantom considering WM, GM, CSF, and bubbles was 24361 mm^3 ($\pm 0.01\%$) more than the sum of all volume measured separately.

EM segmentation was used to segment the image of the brain phantom. Figures 7(a)–7(c) show one slice of the original image, the associated label map produced by EM, and a magnified part of the label, respectively.

As shown in Fig. 7(c), most of the bubbles were detected as GM (gray) because the intensity range of the bubbles was close to the GM intensity range. Using this label map and a module in the 3DSlicer, we measured the volumes of WM (yellow), GM, and CSF (blue). In order to evaluate these measurements, we compared them with the experimental measurements of the compartments. Since EM detected the bubbles as GM, we should compare it (GM_{EM}) by the summation of GM measured experimentally (GM_{exp}) plus the bubbles. Then

$$GM_t = GM_{exp} + \text{bubbles} = 147{,}100 + 24{,}300 = 171{,}400 \text{ mm}^3,$$

where GM_t can be used to estimate the error of the EM segmentation method for calculating the GM. Table 3 shows the EM segmentation method's relative errors. Since GM measured by EM was

$$GM_{EM} = 170{,}678 \text{ mm}^3,$$

then the relative error concerning the total volume is

$$RE1(GM_{EM}) = \frac{|GM_t - GM_{EM}|}{\text{phantom vol}} = \frac{|171{,}400 - 170{,}678|}{901{,}900} \cong 0.0008. \tag{6}$$

However, the relative error concerning GM is

Table 2 Nine samples of different combinations of paraffin wax and SEBS gel, min, max, mean, and standard deviation (StdDev) of measured intensities of them by two scanners 1.5 and 3.0 Tesla. The gray rows are chosen as three tissue to mimic WM, GM, and CSF in the brain.

	Materials			Intensities			
	Sample	SEBS gel (%)	Paraffin (%)	Min	Max	Mean	StdDev
1.5 Tesla	1	100	0	420	585	436	16
	2	90	10	351	419	399	17
	3	80	20	309	350	330	11
	4	70	30	232	271	248	11
	5	60	40	183	201	192	5
	6	50	50	169	179	173	3
	7	40	60	141	165	155	7
	8	30	70	88	140	118	14
	9	20	80	52	87	68	9
3.0 Tesla	1	100	0	404	515	422	18
	2	90	10	344	393	375	13
	3	80	20	235	331	283	30
	4	70	30	193	232	210	11
	5	60	40	172	186	179	4
	6	50	50	160	170	164	3
	7	40	60	138	159	151	5
	8	30	70	82	135	109	13
	9	20	80	51	81	62	7

Table 3 Numbers show the range intensities for three real MRIs acquired from a 3.0-Tesla scanner. GM, WM, and CSF are gray and WMs and cerebrospinal fluid, respectively.

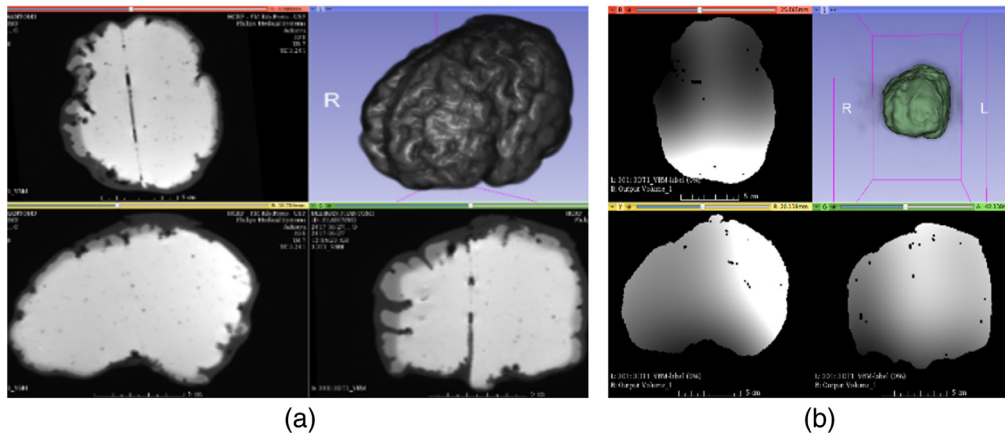|	Tissues	Min	Max	Mean	StdDev
Image 1	GM	0	229	116	19
	WM	0	201	78	32
	CSF	0	85	28	17
Image 2	GM	336	723	484	43
	WM	43	747	289	75
	CSF	0	290	87	42
Image 3	GM	102	156	120	6
	WM	60	118	91	11
	CSF	5	77	42	12

Fig. 6 (a) Complete phantom of the infant's brain with three materials mimics the real brain. N, noise and B, bubbles. (b) The bias effect took place in the brain phantom.
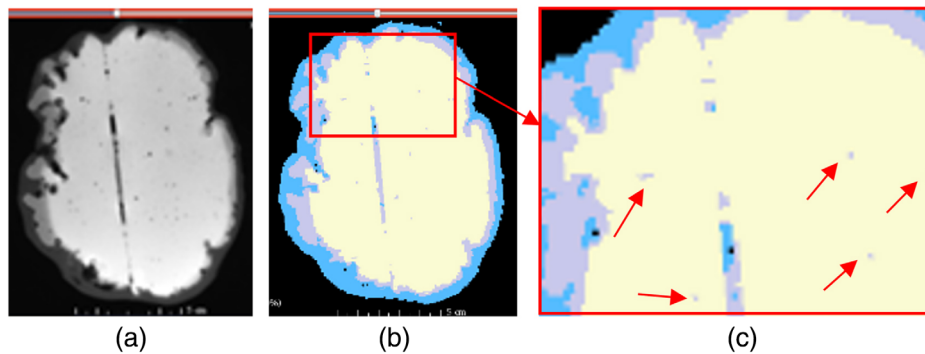

Fig. 7 (a) Original acquired image of brain phantom; (b) label map of phantom by EM-algorithm in 3DSlicer [■ CSF (blue), ■ GM (Gray), and ■ WM (yellow)]; and (c) red arrows show the bubbles, which were detected as GM by EM.

$$\mathrm{RE2(GM_{EM})} = \frac{|\mathrm{GM}_t - \mathrm{GM_{EM}}|}{\mathrm{GM}_t} = \frac{|171{,}400 - 170{,}678|}{171{,}400} \cong 0.004. \tag{7}$$

The relative errors for WM, CSF, and the total volume were calculated in the same way and are presented in Table 4.

Table 4 Relative errors calculated for EM-segmentation. (+b) including bubbles. RE1, relative error concerning the phantom volume and RE2, relative error concerning each compartment.

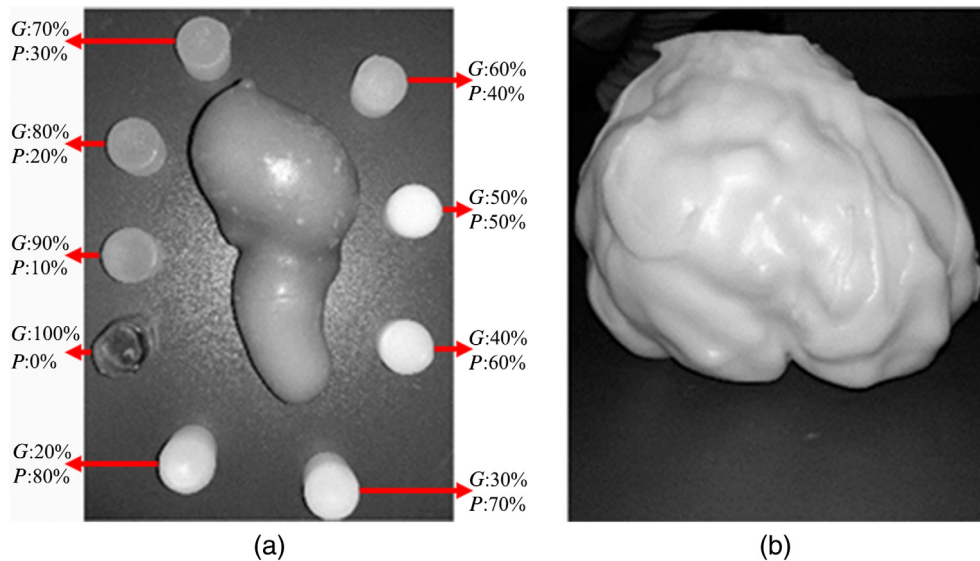|	SEBS gel (%)	Paraffin (%)	Measured volumes in lab (mm³ ±0.03%)	Measured volumes by EM (mm³)	RE1 (±0.03%)	RE2 (±0.03%)
GM	70	30	147,100	170,678 (+b)	0.08	0.4
WM	100	0	531,300	531,021	0.03	0.05
CSF	20	80	199,100	197,920	0.13	0.59
GM + WM + CSF			877,500	—	—	—
Brain phantom			901,800	899,619	0.24	0.24
Bubbles vol			24,300			

Fig. 8 (a) Hypothalamus phantom covered by two samples surrounded by nine different combinations of SEBS gel (G) and paraffin (P). (b) Brain phantom of the infant, diagnosed by Sturge-Weber syndrome, made by three samples ($G = 100\%$, 70%, and 20% with $P = 0\%$, 30%, and 80%, respectively).

The relative error for GM (0.08%) was in good agreement with the error calculated using manual segmentation in Eq. (9), which was 0.07%.

Figure 8(a) shows the photograph of the covered hypothalamus phantom and the nine cylindrical samples. It can be seen that by increasing the percentage of paraffin in the samples, their transparency has decreased. Figure 8(b) shows the photograph of the final brain phantom.

4 Discussion

We propose a new challenging model, i.e., a biomimetic brain phantom with premeasured compartments showing an acceptable accuracy ($\pm 0.03\%$) for quantifying brain volume measurements with MRI. Since previous quantifiers such as manual labeling and simulators have several disadvantages, e.g., time-consuming and limitation in sick brain simulations, this model can be used as an alternative.

Our proposed model was a physical phantom and its image, which provides a quantifier even for technicians to evaluate their manual volume measurement performance. Although there are other methods to evaluate manual volume estimation, e.g., cadavers' muscle dissection[48] for muscle volume estimation, the dissection's accuracy still suffers from human error. It is not always applicable to dissect tissue and measure its volume. Instead, we measured the volume of the phantom using two different methods with acceptable accuracy.

Another advantage of the proposed phantom is the possibility of manipulating the intensity level range of its compartments using a different mixture of SEBS gel and paraffin. This is particularly important because it makes it more difficult for brain volume measurements to estimate volume compartments accurately. In this study, we selected those samples with the intensity level ranges with the highest similarities to the intensities of real brain substances in MRI. One can choose those samples, which have closer intensity ranges, together to be more challenging to distinguish. This can be considered as an advantage of the physical phantoms.

To the best of our knowledge, this was the first physical phantom to mimic a real brain of a human with Sturge-Weber disease. However, we believe it is possible to construct a 3D model for other brain abnormalities and prepare a phantom using our preparation pipeline.

Although undesired bubbles in the phantom were measurable, the bubbles should be avoided in a phantom study. We kept the samples in the oven to eliminate the air bubbles over 5 h, but a few bubbles still appeared in the final phantom when filling the mold. It is left for the future

works to remove all bubbles using another approach. Using injection filling may decrease the bubbles.

5 Conclusions

In this study, a physical brain phantom was constructed using three mixtures of SEBS gel and paraffin (mimicking WM, GM, and CSF in the real brain) and a 3D printed model. Artifacts, i.e., noise, inhomogeneity, and bias effect, appeared in the phantom's acquired image similar to a real brain image. The volumes of the brain phantom compartments were measured experimentally using two methods. The measurements of these volumes and the acquired MR image of the phantom were proposed as a new challenging model to evaluate volume brain measurement methods. To the best of our knowledge, this is the first phantom mimicking a brain with an abnormality. The EM segmentation method was evaluated using the proposed model, and the results showed a correlation between the calculated error using the new metric and the error calculated in the previous study. The estimated relative errors for GM, WM, and CSF were 0.08%, 0.03%, and 0.13% with ±0.03% uncertainty. Although the bubbles in the phantom can be used as a gas-filled lesion that may result from infection, possibly microbleeds, small arteriovenous malformations, or calcifications, the bubbles were an undesired part of this study. For future work, using another approach, for example, injection filling, may help to decrease the bubbles.

Disclosures

The authors declare that they have no conflicts of interest.

Acknowledgments

Ethical approval

All procedures performed in studies involving human participants were in accordance with the local ethics committee, which approved the experimental procedure, and the patient's guardian provided a signed informed consent (CAAE 36460914.4.0000.5440).

References

1. I. Despotovic, B. Goossens, and W. Philips, "MRI segmentation of the human brain: challenges, methods, and applications," *Comput. Math. Methods Med.* **2015**, 450341 (2015).
2. T. Kalincik et al., "Volumetric MRI markers and predictors of disease activity in early multiple sclerosis: a longitudinal cohort study," *PLoS One* **7**(11), e50101 (2012).
3. J. G. Mandell et al., "Volumetric brain analysis in neurosurgery: Part 1. Particle filter segmentation of brain and cerebrospinal fluid growth dynamics from MRI and CT images," *J. Neurosurg. Pediatr.* **15**(2), 113–124 (2015).
4. T. Vân Phan et al., "Evaluation of methods for volumetric analysis of pediatric brain data: the childmetrix pipeline versus adult-based approaches," *NeuroImage* **19**, 734–744 (2018).
5. C. Egger et al., "MRI FLAIR lesion segmentation in multiple sclerosis: does automated segmentation hold up with manual annotation?" *NeuroImage* **13**, 264–270 (2017).
6. M. J. Fartaria et al., "Automated detection of white matter and cortical lesions in early stages of multiple sclerosis," *J. Magn. Reson. Imaging* **43**(6), 1445–1454 (2016).

7. R. Kumar and K. J. Mathai, "Brain tumor segmentation by modified K-mean with morphological operations," *Int. J. Innovative Res. Sci. Eng. Technol.* **6**(8), 16191–16198 (2017).

8. V. Rajinikanth et al., "Segmentation and analysis of brain tumor using tsallis entropy and regularised level set," *Lect. Notes Electr. Eng.* **434**, 313–321 (2018).

9. K. M. Pohl et al., "A hierarchical algorithm for MR brain image parcellation," *IEEE Trans. Med. Imaging* **26**(9), 1201–1212 (2007).

10. N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using K-means clustering algorithm and subtractive clustering algorithm," *Procedia Comput. Sci.* **54**, 764–771 (2015).

11. P. Diniz et al., "Brain tissue segmentation using q-entropy in multiple sclerosis magnetic resonance images," *Braz. J. Med. Biol. Res.* **43**(1), 77–84 (2010).

12. A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," *BMC Med. Imaging* **15**(1), 29 (2015).

13. D. J. Withey and Z. J. Koles, "Medical image segmentation: methods and software," in *Noninvasive Functional Source Imaging Brain and Heart and Int. Conf. Functional Biomedical Imaging, Joint Meeting 6th Int. Symp.*, pp. 140–143 (2007).

14. A. C. Vernon et al., "Effect of chronic antipsychotic treatment on brain structure: a serial magnetic resonance imaging study with ex vivo and postmortem confirmation," *Biol. Psychiatry* **69**(10), 936–944 (2011).

15. B. Aubert-Broche, A. C. Evans, and L. Collins, "A new improved version of the realistic digital brain phantom," *NeuroImage* **32**(1), 138–145 (2006).

16. E. Perilli et al., "A physical phantom for the calibration of three-dimensional x-ray microtomography examination," *J. Microsc.* **222**(2), 124–134 (2006).

17. C. A. Cocosco et al., "Brainweb: online interface to a 3D MRI simulated brain database," *NeuroImage* **5**(no. 4, part 2/4), S425 (2018).

18. S. M. Smith et al., "Advances in functional and structural MR image analysis and implementation as FSL," *NeuroImage* **23**, S208–S219 (2004).

19. M. Azimbagirad et al., "Tsallis-entropy segmentation through MRF and Alzheimer anatomic reference for brain magnetic resonance parcellation," *Magn. Reson. Imaging* **65**, 136–145 (2020).

20. D. L. Collins et al., "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imaging* **17**(3), 463–468 (1998).

21. F. W. Grillo et al., "Patient-specific neurosurgical phantom: assessment of visual quality, accuracy, and scaling effects," *3D Print. Med.* **4**(1), 3 (2018).

22. M. P. Iturralde, *Dictionary and Handbook of Nuclear Medicine and Clinical Imaging*, CRC Press (1989).

23. J. Laing et al., "Patient-specific cardiac phantom for clinical training and preprocedure surgical planning," *J. Med. Imaging* **5**(2), 021222 (2018).

24. B. Bentz et al., "3D printed optical phantoms and deep tissue imaging for in vivo applications including oral surgery," *Proc. SPIE* **10056**, 1005607 (2017).

25. C. Jia et al., "Two-layer heterogeneous breast phantom for photoacoustic imaging," *J. Biomed. Opt.* **22**(10), 106011 (2017).

26. M. Dantuma, R. van Dommelen, and S. Manohar, "A 3D semi-anthropomorphic photoacoustic breast phantom," *Proc. SPIE* **10878**, 108781P (2019).

27. T. L. de Jong et al., "Designing and validating a PVA liver phantom with respiratory motion for needle-based interventions," *Int. J. Comput. Assist. Radiol. Surg.* **14**, 2177–2186 (2019).

28. F. Rengier et al., "3D printing based on imaging data: review of medical applications," *Int. J. Comput. Assist. Radiol. Surg.* **5**(4), 335–341 (2010).

29. A. Chan et al., "Reconstruction and positional accuracy of 3D ultrasound on vertebral phantoms for adolescent idiopathic scoliosis spinal surgery," *Int. J. Comput. Assist. Radiol. Surg.* **14**(3), 427–439 (2019).

30. S. Engelhardt et al., "Flexible and comprehensive patient-specific mitral valve silicone models with chordae tendineae made from 3D-printable molds," *Int. J. Comput. Assist. Radiol. Surg.* **14**(7), 1177–1186 (2019).

31. C. D. Kurth et al., "A dynamic phantom brain model for near-infrared spectroscopy," *Phys. Med. Biol.* **40**(12), 2079 (1995).

32. V. S. Ramachandran, S. Blakeslee, and N. Shah, *Phantoms in the Brain: Probing the Mysteries of the Human Mind*, William Morrow, New York (1998).

33. S. Wood et al., "Design and fabrication of a realistic anthropomorphic heterogeneous head phantom for MR purposes," *PLoS One* **12**(8), e0183168 (2017).

34. A. S. Soliman et al., "A realistic phantom for validating MRI-based synthetic CT images of the human skull," *Med. Phys.* **44**(9), 4687–4694 (2017).

35. M. Mow et al., "Brain perfusion imaging using a reconstruction-of-difference (RoD) approach for cone-beam computed tomography," *Proc. SPIE* **10132**, 1013212 (2017).

36. B. McDermott et al., "Stable tissue-mimicking materials and an anatomically realistic, adjustable head phantom for electrical impedance tomography," *Biomed. Phys. Eng. Express* **4**(1), 015003 (2017).

37. C. Boraxbekk, "Simulating effects of brain atrophy in longitudinal PET imaging with an anthropomorphic brain phantom," *Phys. Med. Biol.* **62**, 5213–5227 (2017).

38. V. Giacometti et al., "Development of a high resolution voxelised head phantom for medical physics applications," *Phys. Medica* **33**, 182–188 (2017).

39. W. Vogt et al., "Phantom-based image quality test methods for photoacoustic imaging systems," *J. Biomed. Opt.* **22**(9), 095002 (2017).

40. A. Pifferi et al., "Mechanically switchable solid inhomogeneous phantom for performance tests in diffuse imaging and spectroscopy," *J. Biomed. Opt.* **20**(12), 121304 (2015).

41. M. O. Culjat et al., "A review of tissue substitutes for ultrasound imaging," *Ultrasound Med. Biol.* **36**(6), 861–873 (2010).

42. L. C. Cabrelli et al., "Copolymer-in-oil tissue-mimicking material with tunable acoustic properties," in *IEEE Int. Ultrasonics Symp.*, pp. 1–4 (2016).

43. L. C. Cabrelli et al., "Stable phantom materials for ultrasound and optical imaging," *Phys. Med. Biol.* **62**(2), 432–447 (2016).

44. Y. Hadadian et al., "Synthesis and characterization of zinc substituted magnetite nanoparticles and their application to magneto-motive ultrasound imaging," *J. Magn. Magn. Mater.* **465**, 33–43 (2018).

45. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. Ser. B* **39**, 1–22 (1977).

46. ANATOMIC, "Cérebro com Artérias em 8 Partes," https://www.anatomic.com.br/produto/cerebro-com-arterias-em-8-partes (2017).

47. S. Pieper et al., "The NA-MIC Kit: ITK, VTK, pipelines, grids and 3D slicer as an open platform for the medical image computing community," in *3rd IEEE Int. Symp. Biomed. Imaging: Nano to Macro*, pp. 698–701 (2006).

48. B. W. Infantolino et al., "The validity of ultrasound estimation of muscle volumes," *J. Appl. Biomech.* **23**(3), 213 (2007).

Mehran Azimbagirad is a postdoc researcher in medical image processing at the University of Western Brittany, France, since 2020. He received his BSc and MSc degrees in applied mathematics and numerical analysis from Razi University and his PhD in applied physics in medicine and biology from University of Sao Paulo, Brazil, in 2019, on brain image segmentation algorithms. Medical image processing, mathematical modelling and machine learning are his main research fields.

Felipe Wilker Grillo is the director of Gphantom LTDA, a company specialized in the development of solutions for medical training using phantoms, mimicking clinical properties of ultrasound, and allowing interaction with medical instruments. He received PhD and MSc degrees in applied physics in medicine and biology from University of Sao Paulo, Brazil.

Yaser Hadadian received his BSc and MSc degrees in solid state physics from Razi University and received his PhD in 2019 from University of Sao Paulo in applied physics in medicine and biology. Currently, he is a postdoctoral researcher at the Intelligent Medical Robotics Laboratory of Gwangju Institute of Science and Technology. His main research field is chemical synthesis and characterization of magnetic nanoparticles for biomedical application, including magneto-motive ultrasound imaging, magnetic hyperthermia, and magnetic particle imaging.

Antonio Adilton Oliveira Carneiro received a PhD in physics applied to medicine and biology from the University of São Paulo in 2001. Currently, he is a full professor at University of São Paulo, coordinator of the Group for Innovation in Biomedical Instrumentation, vice-president of the Brazilian Society of Biomedical Engineering, and member of the board of the Brazilian Societies of Biomedical Engineering, Physics and Medical Physics. His research areas are biomedical instrumentation and theranostics systems by ultrasound, magnetism and optics.

Luiz Otavio Murta Jr. is an associate professor of applied physics and image processing at the University of São Paulo, Department of Computing and Mathematics. He has experience in biomedical engineering, focusing on image-guided therapy, medical image processing, and complex signal processing & analysis, and specifically on image-guided radiation therapy, and cerebral therapy monitoring through image analysis.