Received: 22 June 2021 | Revised: 21 February 2022 | Accepted: 22 February 2022

DOI: 10.1111/jon.12985

JOURNAL OF NEUROIMAGING
THE OFFICIAL JOURNAL OF THE AMERICAN SOCIETY OF NEUROIMAGING
WILEY

CLINICAL INVESTIGATIVE STUDY

Improved neonatal brain MRI segmentation by interpolation of motion corrupted slices

Anouk S. Verschuur[1,2] | Vivian Boswinkel[3,4] | Chantal M.W. Tax[2,5] | Jochen. A.C. van Osch[6] | Ingrid M. Nijholt[1] | Cornelis H. Slump[7] | Linda S. de Vries[8] | Gerda van Wezel-Meijler[3] | Alexander Leemans[2] | Martijn F. Boomsma[1]

[1]Department of Radiology, Isala, Zwolle, The Netherlands

[2]Image Sciences Institute, University Medical Center Utrecht, Utrecht, The Netherlands

[3]Women and Children's Hospital, Isala, Zwolle, The Netherlands

[4]UMC Utrecht Brain Center, Utrecht University, Utrecht, The Netherlands

[5]Cardiff University Brain Research Imaging Centre, Cardiff, UK

[6]Department of Clinical Physics, Isala, Zwolle, The Netherlands

[7]Department of Robotics and Mechatronics, University of Twente, Enschede, The Netherlands

[8]Department of Neonatology, Wilhelmina Children's Hospital, Utrecht, The Netherlands

Correspondence
Anouk S. Verschuur, Department of Radiology (V.3.0), Isala hospital, Dokter van Heesweg 2, 8025 AB Zwolle, The Netherlands.
Email: a.s.verschuur@isala.nl

FUNDING INFORMATION
This research was supported by the Isala Science and Innovation Fund (Isala Hospital, Zwolle, the Netherlands), the Dr. C. J. Vaillant Fund (Landelijke Vereniging van Crematoria, Almere, the Netherlands), and Nutricia Specialized Nutrition (Nutricia Nederland B.V., Zoetermeer, the Netherlands). The sponsors had no role in the design and conduct of the study, collection, management, analysis, and interpretation of the data, or in the preparation, review, approval of the manuscript, or decision to submit the manuscript for publication.

Abstract

Background and Purpose: To apply and evaluate an intensity-based interpolation technique, enabling segmentation of motion-affected neonatal brain MRI.

Methods: Moderate-late preterm infants were enrolled in a prospective cohort study (Brain Imaging in Moderate-late Preterm infants "BIMP-study") between August 2017 and November 2019. T2-weighted MRI was performed around term equivalent age on a 3T MRI. Scans without motion (n = 27 [24%], control group) and with moderate-severe motion (n = 33 [29%]) were included. Motion-affected slices were re-estimated using intensity-based shape-preserving cubic spline interpolation, and automatically segmented in eight structures. Quality of interpolation and segmentation was visually assessed for errors after interpolation. Reliability was tested using interpolated control group scans (18/54 axial slices). Structural similarity index (SSIM) was used to compare T2-weighted scans, and Sørensen-Dice was used to compare segmentation before and after interpolation. Finally, volumes of brain structures of the control group were used assessing sensitivity (absolute mean fraction difference) and bias (confidence interval of mean difference).

Results: Visually, segmentation of 25 scans (22%) with motion artifacts improved with interpolation, while segmentation of eight scans (7%) with adjacent motion-affected slices did not improve. Average SSIM was .895 and Sørensen-Dice coefficients ranged between .87 and .97. Absolute mean fraction difference was ≤0.17 for less than or equal to five

interpolated slices. Confidence intervals revealed a small bias for cortical gray matter (0.14-3.07 cm^3), cerebrospinal fluid (0.39-1.65 cm^3), deep gray matter (0.74-1.01 cm^3), and brainstem volumes (0.07-0.28 cm^3) and a negative bias in white matter volumes (−4.47 to −1.65 cm^3).

Conclusion: According to qualitative and quantitative assessment, intensity-based interpolation reduced the percentage of discarded scans from 29% to 7%.

KEYWORDS
interpolation, motion artifacts, MRI, neonatal, segmentation

INTRODUCTION

Motion artifacts are a common problem in nonsedated neonates undergoing MRI.[1,2] Although measures can be taken to reduce the risk of motion artifacts, such as immobilization and scan time reduction techniques, it cannot be entirely prevented.[3,4] Motion of the neonate during multipacket acquisition can lead to motion artifacts distributed over the scan. While good-quality data are often found between the slices with motion artifacts, segmentation of such scans results in erroneous volume estimations.

Addressing motion artifacts in MRI has been an important research theme. Preprocessing techniques to remove or reduce motion artifacts are computationally intensive and may demand specific acquisition protocols or supplemental data alongside the reconstructed images.[2,5] Examples of these techniques are overlapped slice sampling,[6] specific k-space sampling (e.g., PROPELLER or DISORDER),[7,8] acquisition of multiple image stacks,[9] use of a rapid 3-dimensional scout image,[10] navigation techniques, or use of raw (k-space) data.[2,5] However, these data are often not available when retrospectively processing clinical data.

Recent studies focus on the development of motion correction techniques without the use of supplemental data.[11–14] Most of these studies target adult brain MRI. Translating these techniques to neonatal brain MRI is difficult. In the rapidly developing neonatal brain, maturational changes occur within a short period of time.[15,16] This, and the wide variety of MRI acquisition methods, makes it difficult to create a single algorithm that works for different combinations of maturational stage and acquisition method. To the best of our knowledge, only Khalili et al.[11] developed a toolbox for motion correction in neonatal MRI scans, without the need for supplemental data. They used a deep learning technique, based on MRI scans from neonates at 30 weeks postmenstrual age (PMA).[11] Developing the toolbox involved complex training and testing with a dataset without motion artifacts. Unfortunately, there is no validation of this technique for term equivalent age (TEA; around 40 weeks PMA) MRI scans.[11]

Studies discussing preprocessing-based quality improvement of motion-affected neonatal MRI scans are scarce. We therefore aimed to apply and evaluate a simple and efficient intensity-based interpolation technique without requiring supplemental data, and to inform researchers in the field of neonatal imaging about the opportunities of interpolation methods for retrospective quality improvement. Specifically, we evaluated the performance of a simple, intensity-based, shape-preserving cubic spline interpolation technique to substitute motion-affected slices with interpolated slices in multipacket MRI in neonates scanned at TEA. We then evaluated whether these optimized scans can be used for brain segmentation. This work assumes that only 2-dimensional scans in image space are available with motion-free slices adjacent to motion-affected slices where the degree of motion can be detected either manually or automatically.[17,18]

METHODS

Study population

This study was part of a prospective cohort study, the "BIMP study" (acronym for "Brain Imaging in Moderate-late Preterm infants", The Netherlands trial register; NL6310). Moderate and late preterm infants born between 32^{+0} and 35^{+6} weeks' gestation were enrolled between August 2017 and November 2019 from the neonatal medium, high, or intensive care wards. As part of the BIMP study, these infants underwent an MRI scan around TEA.[19] Infants with congenital anomalies of the nervous system, inborn metabolic problems, congenital infections, chromosomal disorders, or whose parents did not speak sufficient Dutch or English were excluded. In addition, MRI scans showing moderate-severe brain lesions as described by Boswinkel et al.[19] were also excluded from this substudy of the BIMP-study.

Remaining MRI scans were manually scored for severity of motion artifacts that were mostly uniformly distributed over axial slices of the T2-weighted scan (technical medicine researcher, ASV). Three groups were defined: no, mild (<10 slices with minor ringing artifacts), and moderate-severe (>10 slices with minor ringing or ≥1 slice with severe ringing/blurring) motion. Scans without motion artifacts were assigned to the control group. Scans with mild and moderate-severe motion were subdivided into motion artifacts that did (hereinafter referred to as motion group) and did not (excluded for this substudy) visually affect automatic segmentation (radiologist with >10 years of experience [MFB], research-physician [VB], and ASV).

Ethical approval was given by the Central Committee in Research Involving Human Subjects, The Hague, The Netherlands

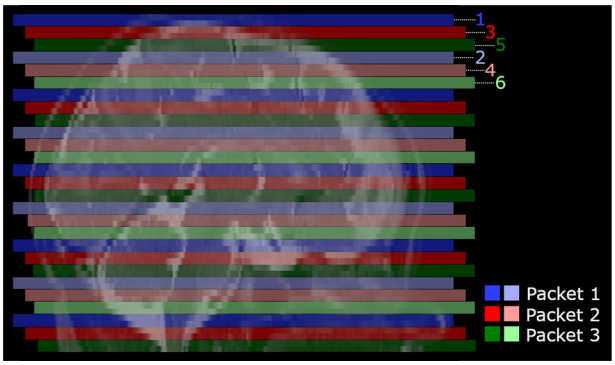

FIGURE 1 An example of interleaved acquisition order with three packages is shown. To minimize interslice cross-talk, slices are first distributed over packages and then the "first odd, then even" rule is applied within each package. Acquisition of sets of slices is as follows: (1) dark blue, (2) light blue, (3) red, (4) pink, (5) dark green, and (6) light green

(NL52323.075.15). Signed informed consent was obtained from both parents for participation in this study.

Data collection

MRI of the brain was performed around 40 weeks PMA using a 3 Tesla MRI scanner (Ingenia, Philips Healthcare, Best, The Netherlands). For this substudy, axial T2-weighted scans were acquired using a turbo spin-echo sequence, sensitivity encoding (in-plane reduction factor 1.7), and multipacket imaging (number of packages: either 2 or 3, with slice scan order for each packet: first uneven then even, as explained in Figure 1); repetition time 5482.8 ms, echo time 110 ms, flip angle 90°, acquired pixel spacing 0.54×0.67 mm^2, reconstructed pixel spacing 0.35×0.35 mm^2, 54 axial slices, and slice thickness 2 mm (no gap). Clinical characteristics including gender, plurality, gestational age, weight and head circumference at birth, and PMA, weight and head circumference at scan were collected from medical charts.

Automatic segmentation

T2-weighted scans were segmented into eight structures (cerebrospinal fluid, cortical gray matter, white matter, deep gray matter, hippocampus, amygdala, brainstem, and cerebellum) using an adapted version of MANTiS (morphologically adaptive neonatal tissue segmentation toolbox).[20,21] MANTiS was developed for 2-dimensional segmentation of neonatal cerebral MRI scans with common anatomical variations in the preterm brain, such as enlarged lateral ventricles. A probability map was used for initial segmentation, followed by morphological filtering to create a subject-specific probability map, which was used for final segmentation.[20] In the adapted version of MANTiS, morphological (watershed) segmentation of cerebrospinal fluid (phase 2) was removed and replaced by thresholding (probability threshold >.9) to improve segmentation in our cohort (with relatively normal brain anatomy).[20,21]

Interpolation

Three intensity-based interpolation methods were considered: linear, spline, and shape-preserving spline interpolation. Interpolation results and difference maps of the three methods were visually assessed in one subject and the structural similarity index (SSIM) was calculated.[22] Uninterpolated slices and background voxels were removed after creating a SSIM-map, to solely assess the impact of interpolation. Eighteen slices, uniformly distributed over the scanned area, were interpolated. The interpolation method with the highest SSIM was used in this study.

Interpolation was performed by removing axial T2-weighted MRI slices and re-estimating them with a 1-dimensional shape-preserving cubic spline interpolation. We used default settings, that is, we used voxel information from four neighboring slices (two before and two after the removed slice) for interpolation.[23,24] This was done in MATLAB 2019a (version 9.6, MathWorks, Natick, Massachusetts, USA).

To assess reliability of interpolation, 18 out of 54 slices per scan (interpolation ratio 1:2) were interpolated in the control group. The 18 interpolated slices were equally distributed over the scan, mimicking the distribution of motion artifacts. More precisely, 11 out of 32 supratentorial (cerebrospinal fluid, cortical gray matter, white matter, deep gray matter, hippocampus, amygdala) slices, three out of eight infratentorial (brainstem and cerebellum) slices, and four out of 14 slices in the transition zone (supratentorial and infratentorial structures present) were interpolated. This did not include field of view boundary slices.

In the motion group, T2-weighted MRI slices that visually showed compromised segmentation were interpolated. The most cranial and caudal slice were not interpolated if motion was present. In both the control and motion group, interpolated scans were re-segmented and structure volumes were calculated.

Visual quality assessment

Quality of interpolation and segmentation (before and after interpolation) was visually assessed by three researchers in consensus (MFB, VB, ASV). Interpolation was assessed by comparing interpolated slices to the same uninterpolated slices (control group) or to adjacent slices (motion group). Areas with severe blurring caused by interpolation were noted for further assessment using segmentation results.

Segmentation was evaluated by overlaying the segmented scan over the T2-weighted scan. Scans were assessed for (1) severe over- or undersegmentation, (2) wrong label assignment, and (3) extracranial segmentation in slices with unreliable interpolation. Improved segmentation in the motion group was defined by removal of severe over- or undersegmentation, no evident areas of wrong label assignment, and no areas with extracranial segmentation. Scans without improvement were excluded for quantitative assessment.

Quantitative assessment

Continuous variables were not normally distributed, so they are summarized with median (interquartile range [IQR]). Clinical characteristics from the motion group and the control group were compared using

the Mann-Whitney U test (for continuous variables) and Fisher's exact test (for dichotomous variables).

Within the control group, SSIM scores comparing scans before and after interpolation were calculated (excluding background voxels). After segmentation, 3-dimensional segmentation overlap between the uninterpolated structures and the interpolated structures was calculated per scan using the Sørensen-Dice coefficient. Dice values of ≥.90 were considered excellent, values between .75 and .89 good, between .5 and .74 moderate, and <.5 poor.

Brain structure volumes were used to assess sensitivity and bias. Sensitivity of interpolation on structure volumes was assessed for supra- and infratentorial slices separately using five subjects from the control group after interpolation of one to five slices. The absolute mean fraction difference was calculated by dividing the absolute difference in volume between interpolated and uninterpolated scans by the uninterpolated volume and calculating the mean between the five subjects.

Structure volumes of the control group before and after interpolation were used to assess potential bias introduced by interpolation. The confidence interval of the mean difference was computed using the paired samples *T*-test with 5000 bootstrapping samples. Structure volumes of the motion group after interpolation and of the control group were used to assess bias between groups. The confidence interval of the mean differences was computed using the independent samples *T*-test with 5000 bootstrapping samples. IBM SPSS statistics (version 26.0, IBM SPSS Statistics for Windows, IBM Corp. Released 2017) was used for statistical analysis. Significance levels were set at $p < .05$.

RESULTS

Study population

Of the 167 infants recruited at birth, in 112 infants T2-weighted scans were eligible for further analysis in this study (Figure 2). Mild ($n = 29$; 26%) and moderate-severe ($n = 56$; 50%) motion artifacts were frequently seen in these T2-weighted scans. Only 27 (24%) scans showed no motion artifacts (control group). Mild motion artifacts did not visually affect segmentation, whereas moderate-severe motion artifacts severely affected segmentation in 33 (29%) scans.

Interpolation

Signal intensity, contrast, and structural changes were visually comparable for linear, spline, and shape-preserving spline interpolation (Figure 3). Spline interpolations resulted in slightly higher tissue contrast as a result of the over- and undershooting nature of spline interpolation. SSIM values were highest for shape-preserving spline interpolation (SSIM = .820).

Quality assessment

At general visual assessment, segmentation of motion-affected slices (Figure 4) improved after interpolation (Figure 5). Differentiation into supra- and infratentorial brain regions presented lower interpolation quality in infratentorial regions compared to supratentorial regions, which is further elaborated below.

Supratentorial regions of the brain

In the control group, interpolated supratentorial T2-weighted slices were radiologically comparable to the uninterpolated slices of the scan. Some contrast reduction and information loss (blurring) were observed as a result of the interpolation process (Figure 6A/a). Segmentation results of these slices (Figure 6B/b) were also visually comparable. In the motion group, interpolation enabled visual structure differentiation in scans with severe motion artifacts (figure 6C/c). Segmentation of these scans improved after interpolation, according to visual assessment (Figure 6D/d).

Infratentorial regions of the brain

In the control group, interpolated infratentorial T2-weighted slices resulted in evident blurring compared to the uninterpolated slices (Figure 7A/a). Subsequently, segmentation of these slices was visually different from uninterpolated slices (Figure 7B/b). In the motion group, T2-weighted scans showed improved tissue differentiation after interpolation (Figure 7C/c). Upon closer inspection, two different cerebellar structures can be recognized in this slice: the inner clearly delineated structure and the outer ghost-like structure. The outlines of these two structures were equal to the outlines of the slice below and above this slice (not visualized). Segmentation showed a decreased volume of cerebellar and brainstem structures after interpolation, compared to the uninterpolated slice (Figure 7D/d).

Transitional zone (both supra- and infratentorial)

Interpolation of inferior MRI slices of the supratentorial cerebrum, which also contains infratentorial structures (brainstem and cerebellum), involves interpolation of superior parts of the cerebellum. Interpolation and segmentation results of slices in the transitional zone were similar for supratentorial structures to the supratentorial results as described above. Interpolation and segmentation results of interpolated infratentorial structures in the transitional zone were comparable to uninterpolated scans as described above.

Overall, improved segmentation was seen in scans from the motion group after interpolation. Scans with motion artifacts in two or more adjacent slices resulted, however, in projection of the artifact of surrounding slices on the interpolated slice. Interpolation and segmen-

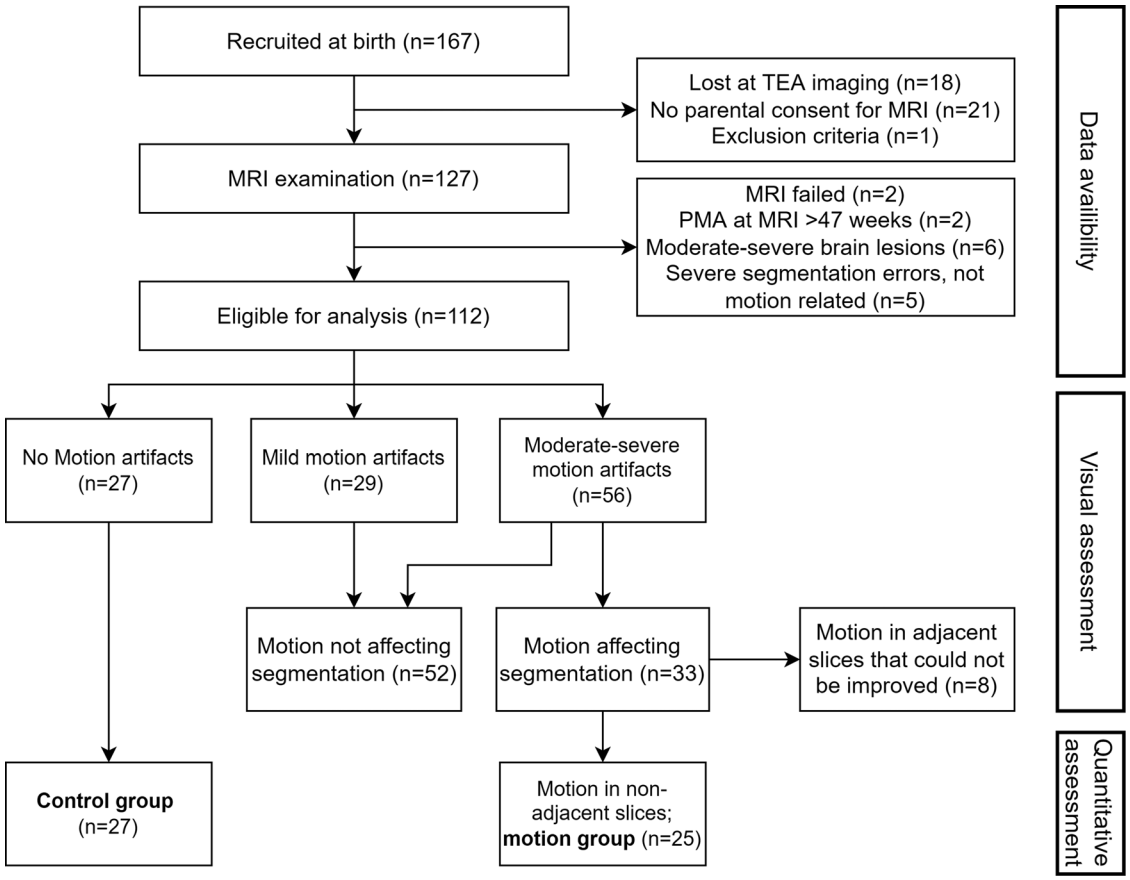

FIGURE 2 Flow diagram of the study population. TEA, term equivalent age; PMA, postmenstrual age; *n*, number of subjects

tation were therefore unsuccessful for eight scans (7%). These scans were excluded for quantitative analysis.

Based on findings from the visual quality assessment, further quantitative analyses were performed using (1) interpolated supra- and infratentorial (including transitional zone) slices from the control group as described in methods, and (2) interpolated supratentorial (including transitional zone) and uninterpolated infratentorial motion-affected slices from the motion group with motion artifacts in nonadjacent slices (n = 25). Infratentorial motion-affected slices were not interpolated because severe ghosting was seen in these slices after interpolation.

Quantitative assessment

Clinical characteristics were not significantly different between the motion and the control group (Table 1).

Within the control group, the average (standard deviation) SSIM score was .895 (0.0078), indicating high similarity between interpolated and uninterpolated scans. In addition, average Sørensen-Dice coefficients of the eight segmented structures were good to excellent (range between .87 and .97; Table 2).

The absolute mean volume fraction difference between interpolated and uninterpolated scans was <0.08 for up to five interpolated slices per scan in the supratentorial brain region. The largest deviation in volume fraction was found for the hippocampus, ranging between 0.04 and 0.07. In all other brain volumes, the absolute mean volume fraction difference was <0.04. For infratentorial interpolation, the absolute mean fraction difference between interpolated and uninterpolated scans was ≤0.17. The lowest sensitivity was found for hippocampal (absolute mean fraction difference 0.06-0.09) and amygdala (absolute mean fraction difference 0.16-0.17) structures. In all other brain volumes, the absolute mean volume fraction difference was <0.05.

Confidence intervals comparing the mean difference between interpolated and uninterpolated structure volumes revealed a potentially small positive bias for cortical gray matter (0.14-3.07 cm^3), cerebrospinal fluid (0.39-1.65 cm^3), deep gray matter (0.74-1.01 cm^3), and brainstem (0.07-0.28 cm^3) volumes and a negative bias in white matter volumes (−4.47 to −1.65 cm^3; Table 3).

Within the motion group, a median of 5 (IQR = 3–10) slices per MRI were interpolated. Comparing the mean difference between the interpolated motion group and the control group does not suggest bias (i.e., confidence intervals include 0; Table 4).

DISCUSSION

In this study, we aimed to apply a simple and efficient intensity-based interpolation technique to overcome motion artifacts in neona-

JOURNAL OF NEUROIMAGING
THE OFFICIAL JOURNAL OF THE AMERICAN SOCIETY OF NEUROIMAGING

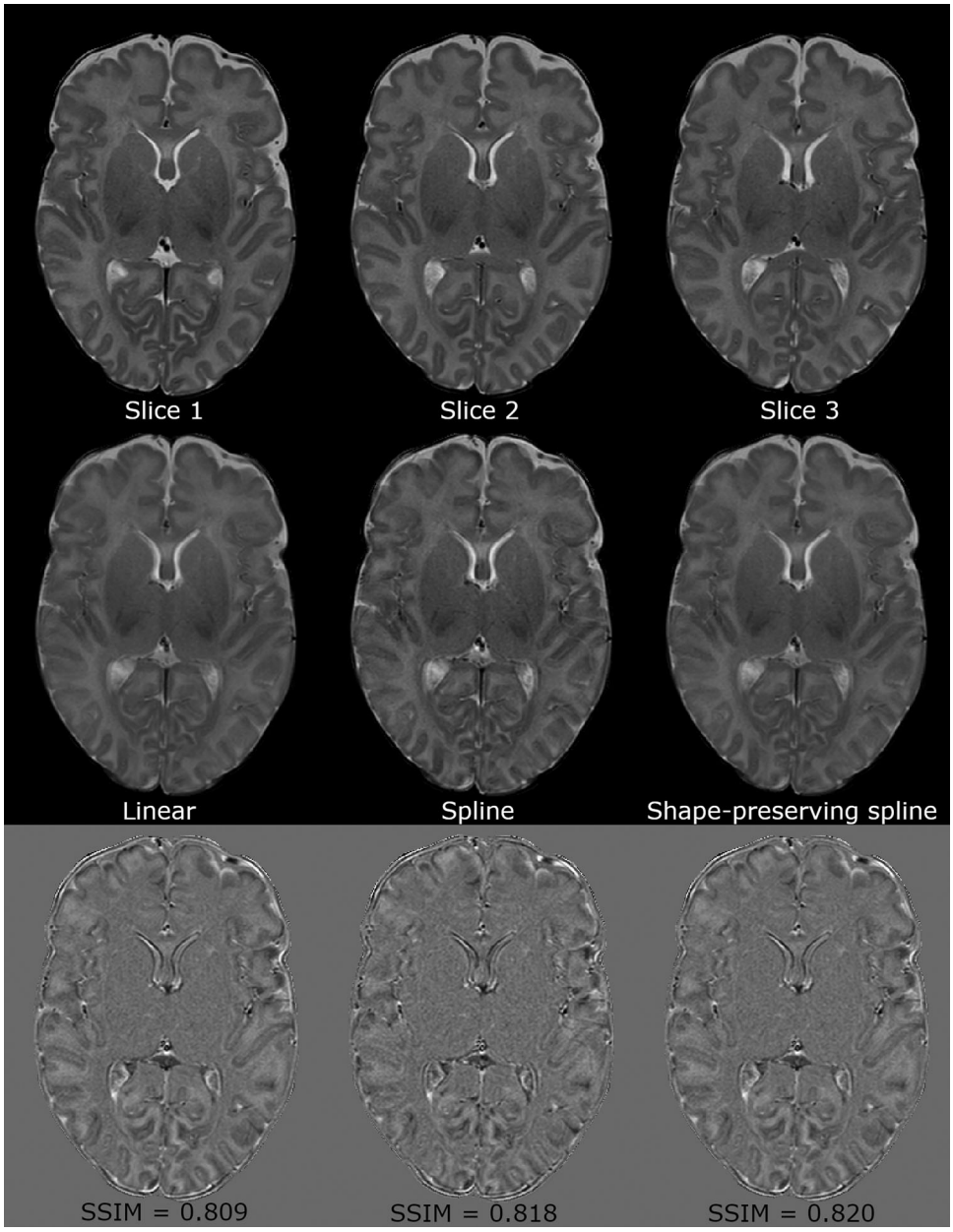

FIGURE 3 Interpolation results with three intensity-based interpolation methods. The top row shows three consecutive slices of an MRI scan. The middle row shows interpolation results of slice 2 using linear (left), spline (middle), and shape-preserving spline (right) interpolation. The bottom row shows the difference maps of the interpolated slices compared to slice 2 in the top row. Structural similarity index (SSIM) is included on the difference maps

tal MRI scans used for automatic brain segmentation. We showed that shape-preserving cubic spline interpolation is a computationally inexpensive method for correction of supratentorial motion artifacts and improves automatic brain segmentation, reducing the percentage of visually unusable scans from 29% to 7% in a homogeneous neonatal population with motion artifacts uniformly distributed over the MRI scan.

Visual assessment demonstrated different results for supratentorial and infratentorial interpolation in both the motion and control group. Infratentorial interpolation and segmentation was inaccurate in most scans. The volumetric change in the axial plane of the cerebellum in the cranial-caudal direction was far larger than for the structures in supratentorial regions. This made interpolation of infratentorial regions, based on adjacent slices above and below, less robust, although not inadequate in all infratentorial slices.

It should be noted that, although careful visual inspection is a good method to assess the performance of interpolation, standardization and quantification is difficult. Therefore, complementary to the subjective visual assessment, SSIM and Sørensen-Dice scores were calculated. The high SSIM score (.895) indicates high similarity between interpolated and uninterpolated scans. This was supported by the good

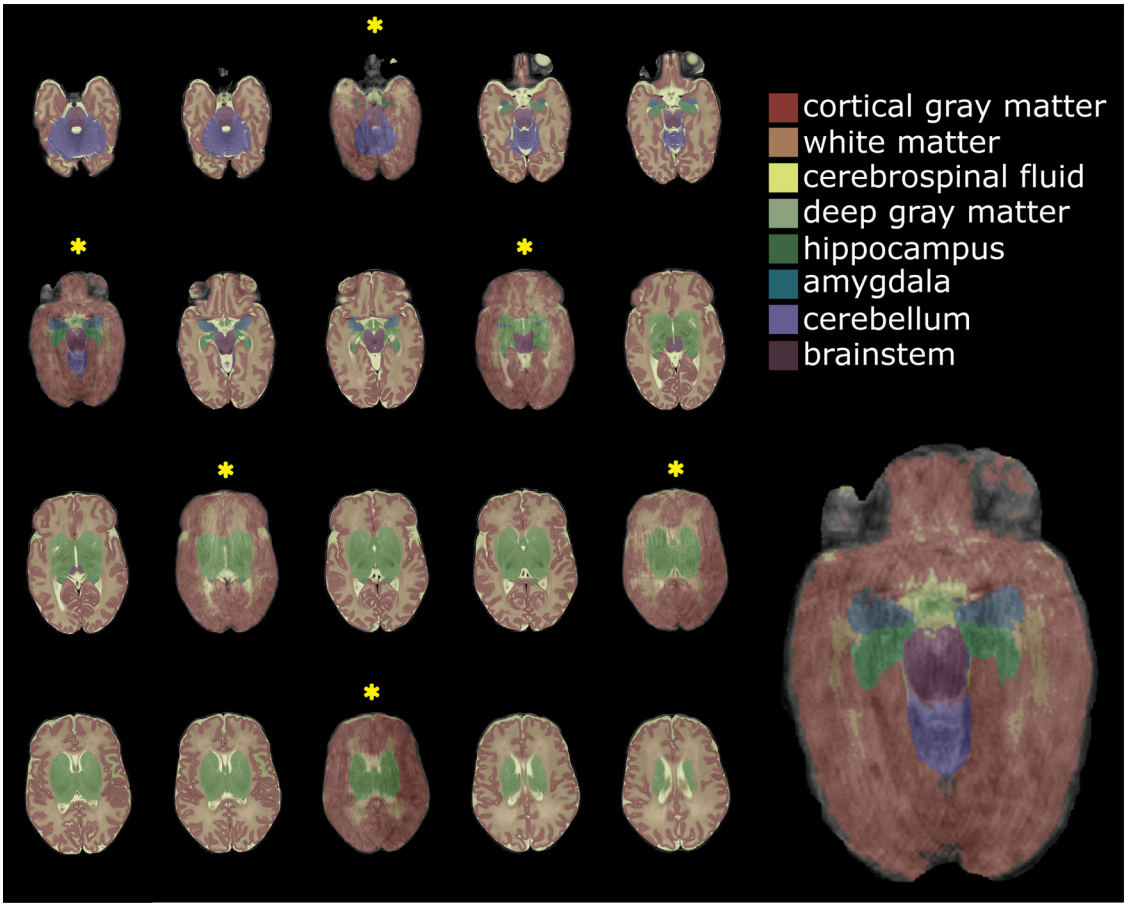

FIGURE 4 Automatically segmented T2-weighted MRI scan (number of packages 3; late preterm infant, gestational age 35^{+2} weeks, postmenstrual age at scan 40^{+1} weeks) with motion artifacts. Slices with motion artifacts are marked with a yellow asterisk and an enlarged example is shown on the right

to excellent Sørensen-Dice coefficients for supratentorial structures (cerebrospinal fluid, cortical gray matter, white matter, deep gray matter, hippocampus, amygdala), indicating that the interpolation technique can properly estimate a removed slice. Excellent Sørensen-Dice coefficients were also found for cerebellar and brainstem regions, although visual inspection showed unreliable interpolation results and blurring. The high Sørensen-Dice coefficients may be a result of comparing two bulk volumes with minimal protrusion/depression, and the small influence from surrounding structures (i.e., less blurring). As a result, segmentation changes mainly took place at the edges of these structures, which has little impact on the Sørensen-Dice coefficient. Moreover, the (visually) less robust interpolation results from individual slices may be neglected in the overall Sørensen-Dice score, due to averaging. Infratentorial interpolation seems therefore feasible, but should be accompanied by careful visual assessment of interpolation and segmentation, since visual results in a single axial slice may be less robust. In general, it should be noted that Sørensen-Dice coefficients are sensitive to region size and shape and tend to be lower for regions with small volumes.[20] Results may therefore not be directly comparable across structures.

A sensitivity analysis showed that small structures, such as the hippocampus and amygdala, are sensitive for interpolation errors. A relatively large part of these structures is re-estimated with interpolation of a single slice. Estimation errors in this slice will result in a relatively large deviation of the original volume, compared to larger brain structures that cover more slices. Over all structures, the absolute mean difference between interpolated and uninterpolated volumes increased mostly, but not exclusively, with increasing number of interpolated slices. The volumetric changes resulting from one interpolation can be positive or negative and therefore counterbalance each other.

Interpolation introduced small bias, as indicated with the confidence interval not including 0 in the control group. However, mean differences in volume before and after interpolation are a small fraction of the median volumes and may only have limited impact on individual volumes. Remarkably, the comparison of the interpolated motion group to the control group did not indicate bias on a group level. This may be a result of the smaller number of interpolated slices in the motion group and of comparing two independent groups, as opposed to the control group where more slices were interpolated and a paired test was used.

JOURNAL OF NEUROIMAGING
THE OFFICIAL JOURNAL OF THE AMERICAN SOCIETY OF NEUROIMAGING

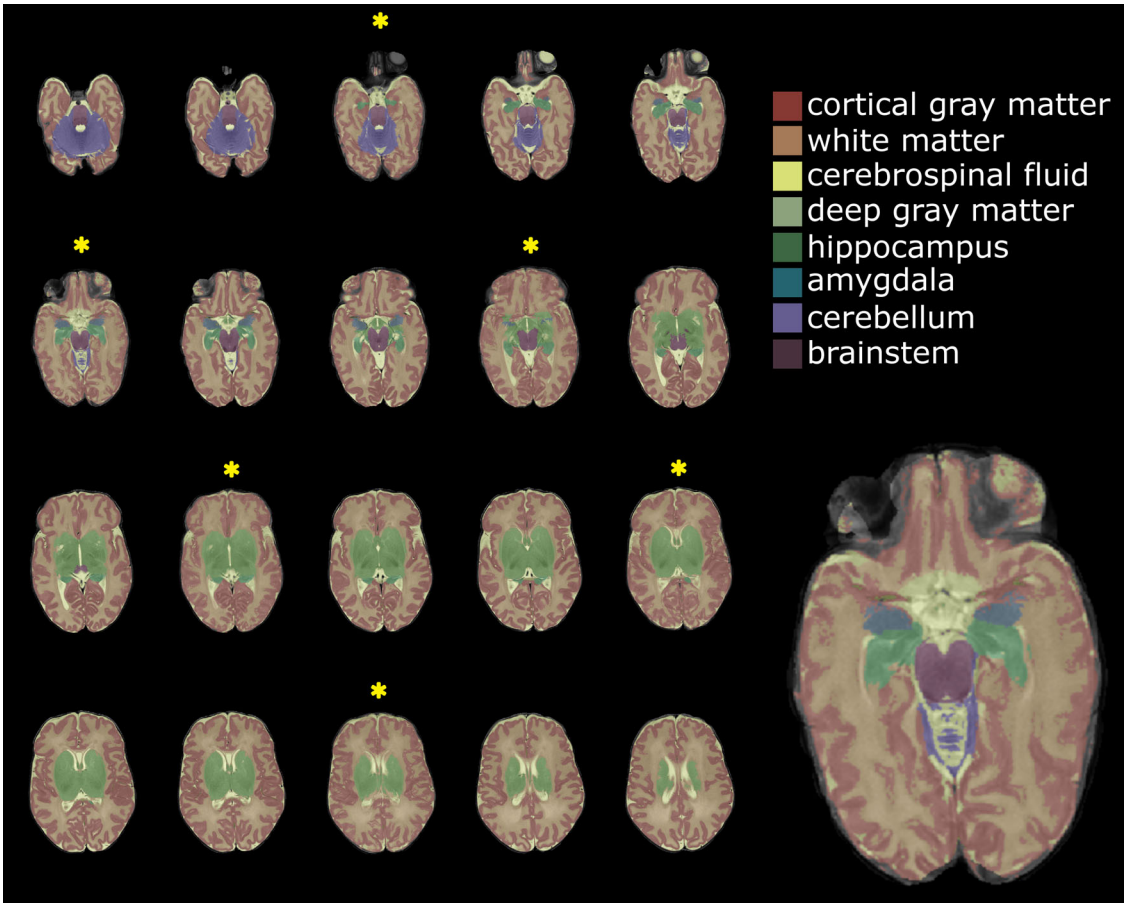

FIGURE 5 Segmentation results after interpolation of supratentorial motion-affected slices—same MRI as in Figure 4. Interpolated slices are marked with a yellow asterisk and an enlarged example (same slice as in Figure 4) is shown on the right

Even though interpolation is preferable to discarding a scan in single-case assessment, this method may introduce a bias when comparing structure volumes between groups. Therefore, this method should only be used with caution in group studies, and with careful assessment of the interpolated scans.

It should be noted that the effects on the presented metrics are primarily a function of the shape, size, and location of the interpolated structures, of the slice thickness, and of the type and extent of the motion artifacts. Therefore, results may only be marginally generalizable and may be different for other data and different methodological choices. Additionally, infratentorial slices of the motion group were not interpolated, which may have affected bias analysis.

Interpolation techniques are broadly used in medical image processing for image resampling, for example, to increase resolution in the slice dimension in MRI scans with anisotropic voxel size.[24–27] With the results of this study, we show that interpolation techniques may also be used to re-estimate motion-corrupted slices without the need for supplemental data. Although the use of more robust correction methods or imaging strategies as discussed in the introduction is preferred, this method enables segmentation and volume computation of a retrospective dataset with motion artifacts distributed over MRI scans, which otherwise would have been excluded from analysis. As in newborn infants MRI is nowadays mostly performed without sedation, this means an important quality improvement and may further contribute to avoiding sedation in this vulnerable population.

Strengths and limitations

Strengths of this study were the use of a computationally efficient method to deal with a complicated problem in neonatal medical imaging. A shape-preserving cubic spline interpolation was chosen to prevent spurious oscillations in areas with sharp intensity changes.[23,24,28] Other strengths were the careful visual assessment of interpolation and segmentation results and the within and between analyses of the subgroups, demonstrating that interpolation is an effective technique to overcome motion artifacts. Moreover, the proposed method does not require additional data, so scan times remain short, which is essential in neonates as they tend to wake-up from their natural sleep and become unsettled.[29]

Although the above-described result are encouraging and are helpful in quantifying volumes that might help identify infants at high risk for developmental delay, these artifact-resolving measures still have some limitations that we would like to elaborate below.

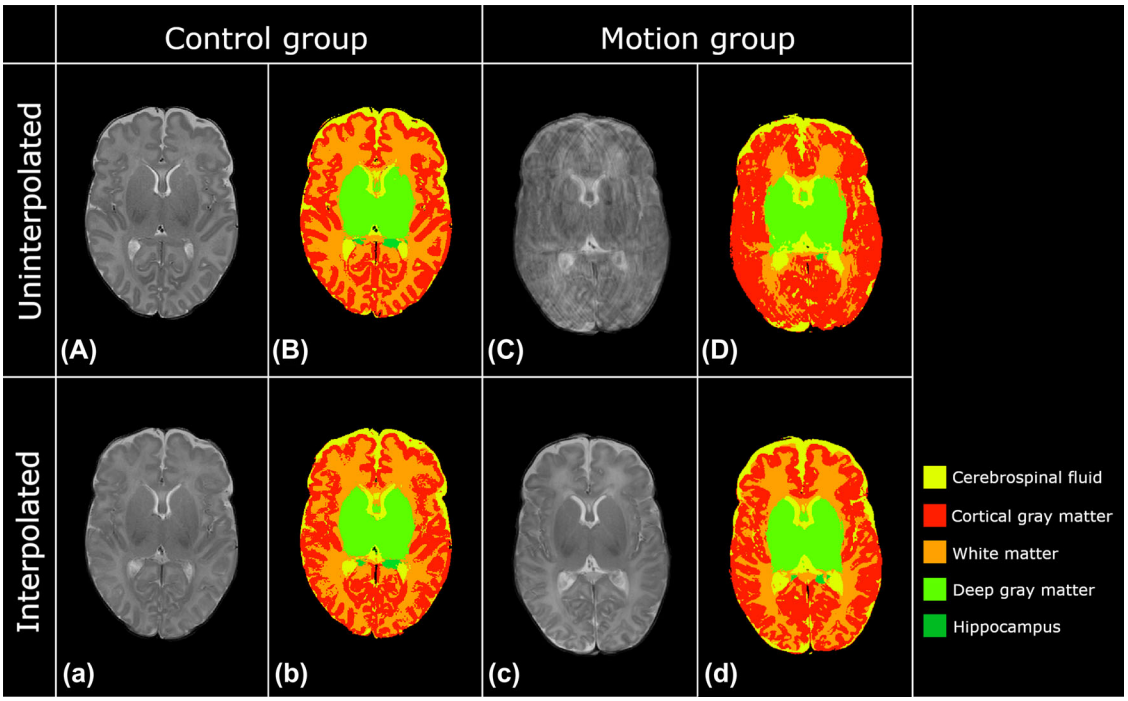

FIGURE 6 Supratentorial interpolation and segmentation results in the control group (moderate preterm infant, gestational age 32^{+5} weeks, postmenstrual age at scan 41^{+3} weeks) and motion group (moderate preterm infant, gestational age 32^{+5} weeks, postmenstrual age at scan 40^{+3} weeks). Figures a/A and b/B show results from the control group; figures c/C and d/D show results from the motion group. Capital letters indicate uninterpolated scans; lower case letters indicate interpolated scans. Improved tissue differentiation is observed after interpolation of motion artifacts

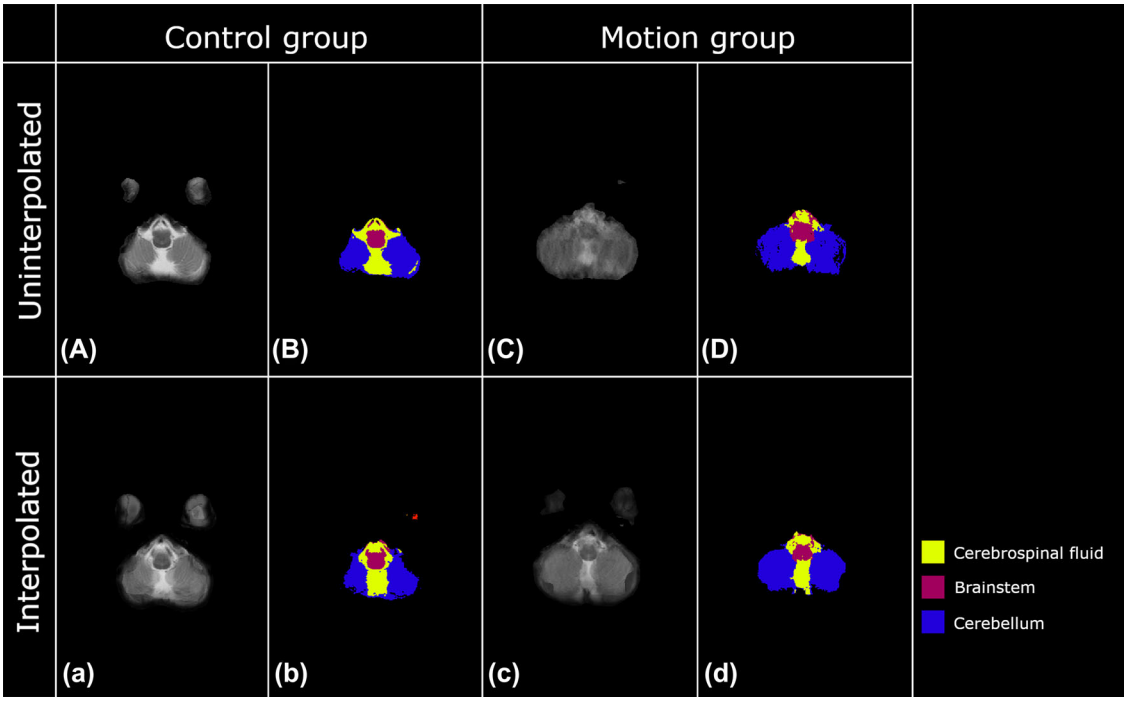

FIGURE 7 Infratentorial interpolation and segmentation results in the control group (late preterm infant, gestational age 35^{+4} weeks, postmenstrual age at scan 40^{+6} weeks) and motion group (late preterm infant, gestational age 35^{+0} weeks, postmenstrual age at scan 45^{+5} weeks). Figures a/A and b/B show results from the control group; figures c/C and d/D show results from the motion group. Capital letters indicate uninterpolated scans; lower case letters indicate interpolated scans. Double boundary issues are observed in the interpolated scans (a and c)

TABLE 1 Participant characteristics

	Control group ($n = 27$)	Motion group ($n = 25$)	p-value
Gestational age in weeks, median (IQR)	33.9 (32.7-35.1)	34.7 (34.3-35.5)	.07
Birth weight, grams, median (IQR)	2275 (2020-2625)	2365 (1962-2750)	.66
Head circumference, cm, median (IQR)	31.5 (30.2-32.5)	32.3 (31.3-33.4)	.15
Boys, *n* (%)	14 (51.9)	12 (48.0)	>.99[a]
Girls, *n* (%)	13 (48.1)	13 (52.0)	
Singleton, *n* (%)	21 (77.8)	19 (76.0)	.44[b]
Twin, *n* (%)	6 (22.2)	4 (16.0)	
Triplet, *n* (%)	0 (0)	2 (8.0)	
Characteristics at term equivalent age MRI			
Postmenstrual age in weeks, median (IQR)	40.9 (40.4-41.6)	40.7 (39.6-42.6)	.98
Weight, g, median (IQR)	3520 (3040-3900)	3440 (3125-3870)	.53
Head circumference, cm, median (IQR)	35.8 (34.2-36.9)	36.0 (34.9-37.4)	.12

Abbreviations: *n*, number of subjects; IQR, interquartile range.
[a]*p*-value was calculated for gender.
[b]*p*-value was calculated for plurality.

TABLE 2 Sørensen-Dice coefficients showing spatial overlap between segmented structures in the control group before and after interpolation

Brain structure	Sørensen-Dice coefficients (SD)
Cortical gray matter	.91 (0.009)
White matter	.91 (0.007)
Cerebrospinal fluid	.87 (0.019)
Deep gray matter	.97 (0.004)
Hippocampus	.89 (0.029)
Amygdala	.87 (0.078)
Cerebellum	.96 (0.005)
Brainstem	.94 (0.008)

Abbreviation: SD, standard deviation.

First, interpolated slices are an estimation of the tissue composition based on adjacent slices. Original slices were discarded before interpolation, meaning that data in the original slices that might contain valuable information were lost. In clinical practice, this technique should be used with caution, since punctate lesions (smaller than the slice thickness) in the original motion artifact scan could potentially be missed in the interpolated scan. We therefore strongly recommend performing a complete visual assessment of the scan, including the slices with motion artifacts before interpolation is applied and use interpolated scans only for segmentation and not for clinical diagnosis.

Second, interpolation of infratentorial slices showed substantial ghosting from structures in slices above and below the interpolated slice, and is therefore not beneficial in quantifying anatomical volumes. This is a major limitation of intensity-based interpolation techniques. Estimation of these isolated bulk tissues may improve with object-based interpolation, so the use of these techniques on neonatal brain scans should be further investigated. Moreover, interpolation using segmentation images may be a more appropriate method for estimating structure volumes in scans with motion artifacts, as grayscale blurring will not affect the final segmentation. Preliminary qualitative results, available at https://github.com/Verschuur95/BIMP_SupplementalData (Part 1), suggest that interpolation of segmentation is feasible. However, its use on neonatal brain scans needs further investigation and optimization.

Third, interpolation of scans with two or more adjacent slices with motion artifacts did not improve segmentation. Interpolation results in either projection of the artifact on the interpolated slice or contrast loss by interpolating multiple adjacent slices at once, making the proposed method unreliable in these more severe cases. Increasing dimensionality of the interpolation function and/or voxel neighborhood might enable slice re-estimation in these scans. This potentially decreases the influence of outliers, but definitely increases computation time. Analyzing these more advanced potentially beneficial techniques was beyond the scope of this study.

Fourth, interpolation results of the motion group were not compared to a ground truth (i.e., an additional T2-weighted scan from the same infant without motion). However, in the motion group, unaffected slices were present between the slices with motion artifacts. Therefore, removing and interpolating slices from the control group reliably simulated the artifact and its correction.

Fifth, it is important to carefully choose segmentation and motion correction techniques as results may be influenced by segmentation algorithm. In this study, we showed that axial slice interpolation enabled 2-dimensional segmentation of motion-affected MRI scans (using MANTiS). However, impact of motion and interpolation on segmentation may differ among segmentation techniques (see https://github.com/Verschuur95/BIMP_SupplementalData [Part 2] for a qualitative comparison). Results suggest that different segmentation techniques are similarly affected by motion artifacts; future work will

TABLE 3 Potential bias introduced by interpolation assessed within the control group using bootstrapped 95% confidence interval

Brain structure $n = 27$	Median (IQR) uninterpolated (cm^3)	Median (IQR) interpolated (cm^3)	Mean difference	95% confidence interval	
Cortical gray matter*	183 (158-212)	179 (159-214)	1.6113	0.14	3.07
White matter*	148 (135-158)	152 (138-162)	−3.0849	−4.47	−1.65
Cerebrospinal fluid*	74.0 (63.1-85.8)	71.2 (62.0-84.9)	0.9975	0.39	1.65
Deep gray matter*	26.4 (24.5-28.4)	25.5 (23.7-27.2)	0.8738	0.74	1.01
Hippocampus	3.31 (2.97-4.16)	3.31 (2.88-4.16)	0.0577	−0.09	0.22
Amygdala	1.39 (1.23-1.71)	1.38 (1.23-1.66)	0.0876	−0.07	0.27
Cerebellum	28.2 (26.0-30.7)	28.2 (25.6-30.9)	0.1474	−0.05	0.34
Brainstem*	7.03 (6.74-7.96)	6.95 (6.64-7.76)	0.1773	0.07	0.28

Note: Confidence interval values including 0 indicate that there is no bias, whereas intervals not including 0 indicate potential bias (denoted with asterisk [*]).
Abbreviations: *n*, number of subjects; IQR, interquartile range.

TABLE 4 Potential bias introduced by interpolation assessed comparing interpolated motion group to the control group using bootstrapped 95% confidence interval

Brain structure $n = 52$	Median (IQR) control group (cm^3)	Median (IQR) motion group (cm^3)	Mean difference	95% confidence interval	
Cortical gray matter	183 (158-212)	197 (175-221)	−15.6	−33.8	2.60
White matter	148 (135-158)	146 (139-160)	−0.18	−7.74	7.43
Cerebrospinal fluid	74.0 (63.1-85.8)	72.8 (51.4-82.6)	2.76	−7.45	12.34
Deep gray matter	26.4 (24.5-28.4)	25.8 (25.1-29.0)	0.23	−1.33	1.84
Hippocampus	3.31 (2.97-4.16)	3.20 (2.78-3.99)	0.20	−0.29	0.70
Amygdala	1.39 (1.23-1.71)	1.48 (1.31-1.75)	−0.07	−0.27	0.15
Cerebellum	28.2 (26.0-30.7)	28.4 (25.8-31.5)	−0.48	−2.39	1.44
Brainstem	7.03 (6.74-7.96)	7.18 (6.75-8.08)	−0.04	−0.48	0.41

Note: Confidence interval values including 0 indicate that there is no bias.
Abbreviations: *n*, number of subjects; IQR, interquartile range.

assess this in a more quantitative way on a larger database and a wider range of segmentation methods. In addition, effects of motion-induced inconsistencies in the slice direction should be further investigated for 3-dimensional segmentation models, as misalignment of slices may result in segmentation errors.

Sixth, by validating the methods with an automatically segmented control group, we were able to solely analyze the effect of interpolation. To analyze accuracy of segmentation after interpolation, one should compare segmentation results before and after interpolation within a neonatal brain phantom with established volumetric dimensions (cerebrospinal fluid, cortical gray matter, white matter, deep gray matter, hippocampus, amygdala, brainstem, and cerebellum), but this was presently not available.

Recommendations

More advanced methods, such as object-based interpolations, may further improve image quality.[30–35] The benefits and reliability of these potentially more robust interpolation methods should be further assessed in neonatal imaging to identify fields of application. Although object-based interpolation may have higher accuracy, in this study we chose a simpler and faster intensity-based interpolation method that requires less computational power.[35]

Moreover, interpolation of probability maps, which are created by using MANTiS segmentation,[20,36] or interpolation of final segmentations is also possible. The potential beneficial effect of using probability maps or segmentation images for interpolation might provide a more reliable estimation of tissue volume, especially in the challenging infratentorial regions.

Finally, the method presented here focuses on the interpolation aspect and relies on prior detection of motion-affected slices, which was done manually in this work. Alternatively, automatic methods could be adopted, which would be a major advantage to make this relatively simple method broadly accessible. A preliminary suggestion for motion detection in data described in this study is available on https://github.com/Verschuur95/Motion_detection.

In conclusion, we showed that relatively simple shape-preserving cubic spline interpolation of T2-weighted MRI scans reduced the percentage of discarded scans (due to motion artifacts) from 29% to 7%. Although the use of more robust methods is preferred when additional data are available, in retrospective data we propose to use this interpolation technique in multipacket MRI in neonates to reduce effects of nonadjacent moderate-severe motion artifacts on segmentation.

ACKNOWLEDGMENTS AND DISCLOSURE

We thank Dr. K. L. Vinken and Dr. H. J. Kuijf (University Medical Centre Utrecht) for their expert advice throughout this project. The authors would like to thank all infants and their parents for their participation. We are grateful to the physicians, (research-) nurses, and radiology technicians involved in the BIMP study for their contribution.

The authors declare no conflict of interest

Verschuur AS, Boswinkel V, van Osch JAC, et al. Cubic interpolation for automatic brain segmentation of MRI motion artifacts in moderate and late preterm infants (Abstract ID: 200144983). Radiological Society of North America 2020.

ORCID

Anouk S. Verschuur https://orcid.org/0000-0001-8737-7078

REFERENCES

1. Arthurs OJ, Edwards A, Austin T, et al. The challenges of neonatal magnetic resonance imaging. Pediatr Radiol 2012;42:1183-94.
2. Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: a complex problem with many partial solutions. J Magn Reson Imaging 2015;42:887-901.
3. Mathur AM, Neil JJ, McKinstry RC, et al. Transport, monitoring, and successful brain MR imaging in unsedated neonates. Pediatr Radiol 2008;38:260-4.
4. Antonov NK, Ruzal-Shapiro CB, Morel KD, et al. Feed and wrap MRI technique in infants. Clin Pediatr 2017;56:1095-103.
5. Godenschweger F, Kägebein U, Stucht D, et al. Motion correction in MRI of the brain. Phys Med Biol 2016;61:R32-56.
6. Cordero-Grande L, Hughes EJ, Hutter J, et al. Three-dimensional motion corrected sensitivity encoding reconstruction for multi-shot multi-slice MRI: application to neonatal brain imaging. Magn Reson Med 2018;79:1365-76.
7. Pipe JG. Motion correction with PROPELLER MRI: application to head motion and free-breathing cardiac imaging. Magn Reson Med 1999;42
8. Cordero-Grande L, Ferrazzi G, Teixeira RPAG, et al. Motion-corrected MRI with DISORDER: distributed and incoherent sample orders for reconstruction deblurring using encoding redundancy. Magn Reson Med 2020;84:713-26.
9. Kuklisova-Murgasova M, Quaghebeur G, Rutherford MA, et al. Reconstruction of fetal brain MRI with intensity matching and complete outlier removal. Med Image Anal 2012;16:1550-64.
10. Polak D, Splitthoff DN, Clifford B, et al. Scout accelerated motion estimation and reduction (SAMER). Magn Reson Med 2022;87:163-78.
11. Khalili N, Turk E, Zreik M, et al. Generative adversarial network for segmentation of motion affected neonatal brain MRI. Presented at the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention; October 13–17, 2019; Shenzen.
12. Duffy BA, Zhang W, Tang H, et al. Retrospective correction of motion artifact affected structural MRI images using deep learning of simulated motion. Presented at the 1ts Conference on Medical Imaging with Deep Learning (MIDL); July 4–6, 2018; Amsterdam.
13. Küstner T, Armanious K, Yang J, et al. Retrospective correction of motion-affected MR images using deep learning frameworks. Magn Reson Med 2019;82:1527-40.
14. Sommer K, Saalbach A, Brosch T, et al. Correction of motion artifacts using a multiscale fully convolutional neural network. AJNR Am J Neuroradiol 2020;41:416-23.
15. Ortinau C, Neil J. The neuroanatomy of prematurity: normal brain development and the impact of preterm birth. Clin Anat 2015;28:168-83.
16. Stiles J, Jernigan TL. The basics of brain development. Neuropsychol Rev 2010;20:327-48.
17. Liu Z, Wang Y, Gerig G, et al. Quality control of diffusion weighted images. Presented at the SPIE Medical Imaging; February 13–18, 2010; San Diego.
18. Zhou Z, Liu W, Cui J, et al. Automated artifact detection and removal for improved tensor estimation in motion-corrupted DTI data sets using the combination of local binary patterns and 2D partial least squares. Magn Reson Imaging 2011;29:230-42.
19. Boswinkel V, Krüse-Ruijter MF, Nijboer - Oosterveld J, et al. Incidence of brain lesions in moderate-late preterm infants assessed by cranial ultrasound and MRI: the BIMP-study. Eur J Radiol 2021;136:109500.
20. Beare RJ, Chen J, Kelly CE, et al. Neonatal brain tissue classification with morphological adaptation and unified segmentation. Front Neuroinform 2016;10:12.
21. Morphologically adaptive neonate tissue segmentation toolbox – adapted. https://github.com/DevelopmentalImagingMCRI/mantis/tree/BIMP-NoLargeVentricles2. Accessed 9 Sept 2020.
22. Wang Z, Bovik AC, Sheikh HR, et al. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004;13:600-12.
23. Wolberg G, Alfy I. Monotonic cubic spline interpolation. Presented at the Proceedings of the Computer Graphics International Conference; June 7–11, 1999; Alberta.
24. Fritsch FN, Carlson RE. Monotone piecewise cubic interpolation. SIAM J Numer Anal 1980;17:238-46.
25. Lehmann TM, Gonner C, Spitzer K. Survey: interpolation methods in medical image processing. IEEE Trans Med Imaging 1999;18:1049-75.
26. Neubert A, Salvado O, Acosta O, et al. Constrained reverse diffusion for thick slice interpolation of 3D volumetric MRI images. Comput Med Imaging Graph 2011;36:130-8.
27. Ravichandran CG, Ravindran G. Inter-slice reconstruction of MRI image using one dimensional signal interpolation. Int J Comput Sci Netw Secur 2008;8:351-6.
28. Steffen M. A simple method for monotonic interpolation in one dimension. Astron Astrophys 1990;239:443-50.
29. Dagia C, Ditchfield M. 3T MRI in paediatrics: challenges and clinical applications. Eur J Radiol 2008;68:309-19.
30. Goshtasby A, Turner DA, Ackerman LV. Matching of tomographic slices for interpolation. TMI 1992;11:507-16.
31. Higgins WE, Orlick CJ, Ledell BE. Nonlinear filtering approach to 3-D gray-scale image interpolation. IEEE Trans Med Imaging 1996;15:580-7.
32. Grevera GJ, Udupa JK. Shape-based interpolation of multidimensional grey-level images. IEEE Trans Med Imaging 1996;15:881-92.
33. Lee T, Lin C. Feature-guided shape-based image interpolation. IEEE Trans Med Imaging 2002;21:1479-89.

34. Lee T, Wang W. Morphology-based three-dimensional interpolation. IEEE Trans Med Imaging 2000;19:711-21.
35. Leng J, Xu G, Zhang Y. Medical image interpolation based on multi-resolution registration. Comput Math Appl 2013;66:1-18.
36. Ashburner J, Friston KJ. Unified segmentation. Neuroimage 2005;26:839-51.

How to cite this article: Verschuur AS, Boswinkel V, Tax CMW, et al. Improved neonatal brain MRI segmentation by interpolation of motion corrupted slices. Journal of Neuroimaging. 2022;32:480–492. https://doi.org/10.1111/jon.12985