CLINICAL RESEARCH ARTICLE

Comparison of MRI and neurosonogram 1- and 2-dimensional morphological measurements of the newborn corpus callosum

Michael Mills[1], Vincent Pelling[1], Lisa M. Harris[1], Joely Smith[2], Neil Aiton[3,4], Heike Rabe[3,4] and Jose Ramon Fernandez-Alvarez[3,4]

BACKGROUND: Developmental abnormalities of the corpus callosum (CC) are linked to multiple neuro-developmental disorders, for which neonatal neuroimaging may allow earlier diagnosis and intervention. MRI is often considered the most sensitive imaging modality to white matter changes, while neurosonogram (NS) remains the clinical staple. This study assesses the correlation between MRI and US measurements of the neonatal CC using a protocol derived from established methodologies.

METHODS: MR and NS images from an existing cohort of term infants (≥37 weeks gestational age) were studied. Length and area measurements of the CC made with linear (LUS) and phased array US (PUS) data were compared to those from MRI. Intra-observer reliabilities were estimated.

RESULTS: Moderate-to-strong correlation strengths were observed for length measurements and the total area of the CC. Sectional area measurements showed poorer correlations. Bland–Altman plots support improved correspondence of length and total area measurements. LUS data appeared to correspond closer to MRI. All three modalities showed comparable repeatability.

CONCLUSION: NS correlates well with some MRI measurements of the CC and shows similar levels of repeatability, making them possibly interchangeable. Use of LUS, a technique rarely used for NS, may be preferable to the standard approach for morphological studies.

INTRODUCTION

The corpus callosum (CC) is a bundle of hundreds of millions of myelinated and non-myelinated nerve fibres connecting the two hemispheres of the brain.[1] It has been shown to be involved in functions such as sound localisation[2] and stereoscopic vision,[3] which has led to the hypothesis that its evolution occurred to facilitate integration of sensory signals between the brain's two hemispheres.[4,5]

Developmental abnormalities of the CC are linked to multiple neuro-developmental disorders, such as attention deficit hyperactivity disorder, autism, dyslexia and foetal alcohol spectrum disorder.[6–8] Early detection of the malformation or underdevelopment of the CC, which are associated with these disorders, allows for earlier interventions on the part of health-care professionals and local public sector providers. Although contradictory results can be found in the literature, outcomes are often improved as a result of intervention programmes for children with known neurological disorders or learning disabilities.[9–11]

Foetal and neonatal neuroimaging using brain magnetic resonance imaging (MRI) or neurosonogram (NS), including the CC, is in common clinical practice.[12] NS is often considered the preferred imaging modality for studies of the neonatal brain as it can be performed at the bedside, is non-ionising, requires little disturbance of the neonate and is relatively inexpensive compared to other imaging modalities.[13] MRI, however, is often considered to be the most sensitive imaging modality to changes in white matter structures and measurements of brain growth[14–16] and has been used in many studies that correlate measurements of the CC with neurological outcomes. MRI is also not limited to specific imaging orientations and locations, while NS may only be performed through specific viewing windows: the neonatal fontanelles. However, MRI is time consuming, expensive, often requires patient sedation and is acoustically noisy.

Recent publications show no significant differences on comparison of linear measurements of some neuroanatomic structures when obtained using MR and ultrasound (US): width, height, length and depth of the corpus callosum, cerebellum, pons, cavum septum pellucidum, fourth ventricle, basal ganglia, and caudate head.[17] Measurements of the ventricular index and diameter of the posterior horn of the lateral ventricle, extracerebral space, inter-hemispheric fissure and cortex of the cingulate gyrus did show significant differences.[17] Strong correlations have also been reported between ventricular volumes for paediatric patients as measured on MR and US images.[18]

In this study, we aimed to investigate the level of correspondence between measurements of the neonatal CC as measured from MRI and US data sets using phased array US (PUS) and linear array US (LUS) transducers and established methodologies for these techniques in a group of neonates undergoing cooling therapy after a diagnosis of hypoxic ischemic encephalopathy (HIE). We also aimed to assess the intra-observer reliability of these measurements.

MATERIALS AND METHODS

We performed a retrospective cohort study of 39 term neonates, diagnosed with HIE, admitted to a tertiary surgical neonatal department from January 2014 to November 2015, undergoing

[1]Department of Radiological Sciences, Brighton & Sussex University Hospitals NHS Trust, Brighton, UK; [2]Department of Imaging, Imperial College Healthcare NHS Trust, London, UK; [3]Brighton and Sussex Medical School, Brighton, UK and [4]Department of Neonatology, Brighton & Sussex University Hospitals NHS Trust, Brighton, UK
Correspondence: Michael Mills (Michael.Mills2@nhs.net)

356

total body cooling therapy and imaged as part of the normal clinical pathway.

All imaging was performed on a 1.5T Siemens Symphony MRI scanner (Siemens, Erlangen, Germany), and a Philips HD11 XE US scanner (Philips, Eindhoven, The Netherlands). US images were obtained with either a 4–12 MHz PUS transducer or a 5–12 MHz LUS transducer, using presets designed for neonatal cranial use. MR images were acquired using the standard clinical protocol, which included the T1-weighted sagittal and T2-weighted turbo inversion recovery magnitude (TIRM) coronal images used within this study. Slight variations in imaging parameters occurred across patients. Typical sequence parameters include a repetition time (TR) = 450 ms and echo time (TE) = 8.6 ms for the T1-weighted sequences, while for the T2 TIRM sequence TR = 8850 ms, TE = 129 ms and inversion time (TI) = 2500 ms were typical. T1- and T2-weighted images were generally acquired with a voxel size of $0.45 \times 0.45 \times 4.00$ mm.

Multiple data sets were available for US images and MRI. To minimise selection bias, only images from data sets acquired within a 7-day period between all three imaging sessions were selected for analysis. Data sets were then chosen based on the identification of anatomical landmarks and ensuring the image captured the entire extent of the CC in the midsagittal plane. For MR, the midsagittal slice was identified as that which clearly showed the patient's nose and for which the cerebellar vermis was visualised and the spinal cord well resolved. For US, the superior border of the CC had to be clearly identifiable against its surroundings and the cavum septum pellucidum, the cerebellar vermis and the aqueduct had to all be visible on one image. Coronal slices on which to assess the thickness of the corpus callosum were initially selected based on a clear visualisation of the interhemispheric fissure, cavum septum pellucidum and the foramina of Monro. Figure 1 shows, for a single patient, example sagittal images.
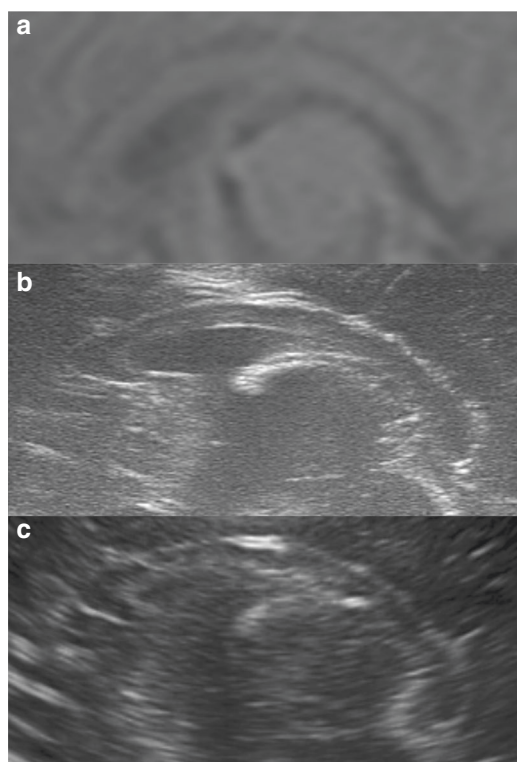

Fig. 1 Example midsagittal images of a single patient from magnetic resonance imaging (**a**), linear array ultrasound (US) (**b**) and phased array US (**c**)

All measurements on NS and MRI were performed as described previously for newborn infants by a single reviewer (M.M.), a medical physicist with experience of viewing medical images, under the guidance of a Neonatologist with expertise in neuroimaging (J.R.F.-A.).[17,19,20] Four lengths and eight areas were determined on the midsagittal images, while a single length was measured on the coronal images as follows. The long axis diameter of the CC was defined as the line connecting the anterior edge of the genu to the posterior edge of the splenium. From the long axis line, all sectional areas of the CC were then defined. Seven regions were defined by drawing perpendicular lines to the long axis midpoint, at one third and at one fifth from the posterior end and at one third from the anterior end. A final line was drawn perpendicular to the long axis line at the inner border of the anterior notch of the CC. These divisions defined the rostrum, genu, rostral body, anterior midbody, posterior midbody, isthmus and splenium as defined by Witelson (Figure 2).[20] The body width of the CC was defined as the thickness of the CC at its midpoint, perpendicular to the long axis. The genu width and splenium width were then defined as the greatest diameter when drawing a perpendicular line from the most anterior edge to the most posterior edge of the boundaries of the genu or splenium sections (as defined above). The area of each section of the CC was summed to yield its total area. Figure 2 shows a schematic of the sagittal measurements and Figure 3 shows this process in four stages on an example anonymised US image from one of the included subjects. A single length measurement in the coronal plane, the coronal width (described by Liu et al. as the antero-posterior diameter), was also obtained by measuring the thickness of the CC above the cavum septum pellucidum and below the inter-hemispheric fissure, perpendicular to its edges.[19]

Prior to all measurements, a histogram normalisation of the image window and level was performed. This was performed to improve the consistency of viewing conditions. Measurements were then obtained using the length calliper and freehand area measurement tools of the local PACS viewer (Philips IntelliSpace PACS 4.4 Enterprise, Eindhoven, The Netherlands). Image zooming by 400% was also performed to ensure the CC encompassed as much of the viewing monitor as possible. No additional image processing occurred.

Intra-observer variability was assessed by performing ten repeat measurements on three randomly selected neonate's images, for which LUS and PUS transducer and MRI data were present. The reviewer was blinded to previous results. Intra-observer error was
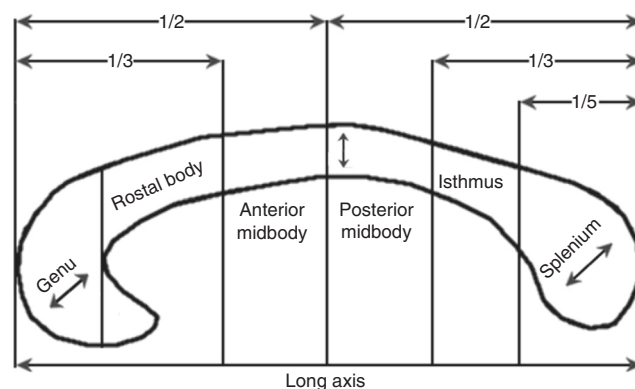

Fig. 2 Schematic of the corpus callosum showing the divisions of the Witelson scheme. Clockwise these regions correspond to the rostrum, genu, rostral body, anterior midbody, posterior midbody, isthmus and splenium. The total area is the sum of these seven areas. The four sagittal length measurements are also displayed. Adapted from Witelson.[20]
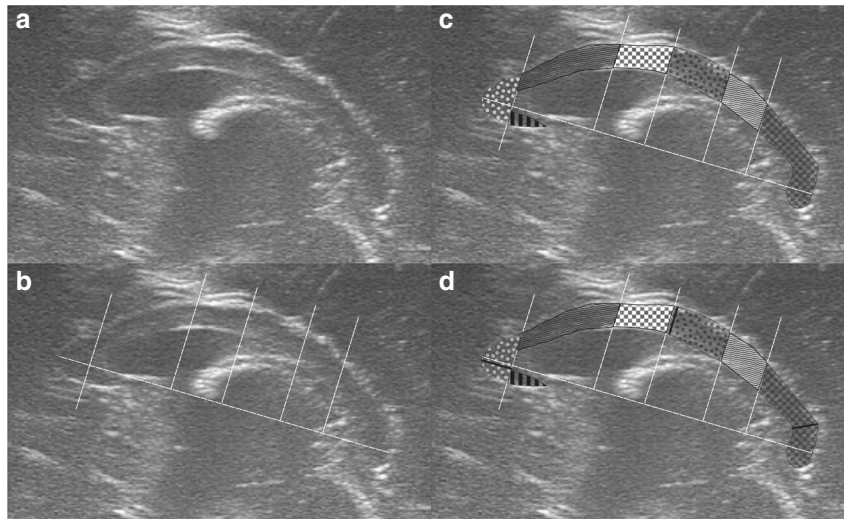
Comparison of MRI and neurosonogram 1- and 2-dimensional morphological...
M Mills et al.

357

Fig. 3 Example linear array US image of the corpus callosum (CC) demonstrating the measurement procedure for measurements obtained in the midsaggital plane. The midsaggital image is first identified (**a**). The longest axial diameter is then defined as a line from the most anterior to the most posterior boundary of the corpus callosum. The seven regions outlined in Figure 2 are defined perpendicular to this line (**b**). These perpendicular bisectors of the long axis are defined by drawing perpendicular lines to the long axis midpoint, at one third and at one fifth from the posterior end and at one third from the anterior end. A final line is drawn perpendicular to the long axis line at the inner border of the anterior notch of the CC. The seven regions of the CC are then drawn manually (**c**). The body width is defined as the thickness of the CC at its midpoint, perpendicular to the long axis. The genu width and splenium width are then defined as the greatest diameter when drawing a perpendicular line from the most anterior edge to the most posterior edge of the boundaries of the genu or splenium sections (**d**)

estimated using the coefficient of variation (COV):

$$COV = \frac{\sigma}{\bar{x}} \times 100\%$$

where σ and $\bar{x}$ are the standard deviation and mean recorded values. The results across the three patients' images were then pooled.

All analysis was performed using SPSS23 (IBM, Armonk, NY). Descriptive statistics, median measurement and interquartile range (IQR), and correlation strengths from comparisons of measurements from different modalities are shown in Table 1. Statistical analysis of correlations was performed using the Spearman's correlation coefficient, assuming non-normal data distribution: P value <0.001 was taken as statistically significant. Correlation by Spearman's coefficient was graded: strong ($\rho > 0.7$), moderate ($0.4 \geq \rho \leq 0.7$), weak ($\rho < 0.4$). In addition, Bland–Altman plots were constructed for each measurement performed comparing the MRI to the US technique. Sets of 95% confidence intervals were constructed for these plots as:

$$CI_{\pm} = \bar{d} \pm 1.96s$$

where $\bar{d}$ is the mean of the calculated differences between imaging methods (for example, the difference in measurements of the long axis length as measured on MRI and LUS images) and s the standard deviation of the differences. Institutional Review Board approval was not required as this was a retrospective analysis of anonymised standard patient information collected during routine care. Clinical Governance approval, as per local hospital guidance, was granted.

Table 1. Corpus callosum measurements

$N = 32$	MRI (IQR)	PUS (IQR)	LUS (IQR)	MRI-PUS, ρ	MRI-LUS, ρ
Long axis	43.75 (4.93)	45.00 (3.25)	44.30 (1.60)	0.72[a]	0.87[a]
Body	2.30 (1.20)	2.20 (0.70)	2.40 (0.90)	0.42	0.44
Genu	4.45 (2.55)	4.10 (1.45)	4.80 (2.00)	0.69[a]	0.64[a]
Splenium	4.65 (1.15)	5.60 (2.45)	5.40 (1.90)	0.49	0.71[a]
Area 1	0.04 (0.04)	0.06 (0.08)	0.07 (0.05)	0.52	0.47
Area 2	0.23 (0.21)	0.19 (0.18)	0.25 (0.12)	0.65[a]	0.79[a]
Area 3	0.29 (0.11)	0.25 (0.08)	0.25 (0.06)	0.30	0.35
Area 4	0.20 (0.06)	0.15 (0.04)	0.19 (0.05)	0.47	0.06
Area 5	0.16 (0.06)	0.15 (0.06)	0.16 (0.05)	0.36	0.25
Area 6	0.16 (0.07)	0.14 (0.06)	0.13 (0.06)	0.25	0.39
Area 7	0.44 (0.13)	0.47 (0.21)	0.47 (0.16)	0.68[a]	0.65[a]
Total area	1.48 (0.53)	1.42 (0.44)	1.57 (0.56)	0.59[a]	0.67[a]
Coronal width	1.95 (0.40)	2.00 (0.40)	2.10 (0.50)	0.57	0.75[a]

Magentic Resonance Imaging (MRI), phased array ultrasound (PUS) and linear array ultrasound (LUS) measurements displayed as median and interquartile range (IQR) in mm for linear measurements and cm^2 for area. Spearman's Rho (ρ) correlation coefficients between MRI and the two US modalities are also displayed. A P value <0.001 is considered statistically significant. Areas 1–7 indicate the regions shown in Figure 2, starting with the rostrum and ending with the splenium
[a]Denotes a significant correlation

RESULTS
Thirty two neonates who underwent MRI and US provided useable data: seven were excluded owing to missing data or the entire CC not residing within the image. All included studies were of term babies (53% male), with a median birth weight of 3415 g (IQR 762 g).

Statistically significant correlations (all moderate-to-strong) were observed for the long axis, genu and sectional areas 2 and 7 as well as for the total area, for both modality comparisons. The LUS to MRI comparison also yielded significant correlations (strong) for the splenium length and coronal width measurements. Moderate correlation strengths were seen for all the other

358

Table 2. Pooled intra-observer coefficient of variation

	MRI	PUS	LUS
Long axis	1.3	1.6	1.2
Body width	16.6	13.3	9.5
Genu width	12.7	12.3	9.9
Splenium width	11.8	12.0	8.9
Area 1	19.9	15.0	17.6
Area 2	18.1	17.9	15.2
Area 3	13.8	8.9	7.1
Area 4	16.5	10.4	9.2
Area 5	13.5	12.8	10.4
Area 6	13.0	12.0	11.0
Area 7	8.5	7.5	9.9
Total area	5.9	4.4	4.7
Coronal width	13.7	10.2	10.1

Ten repeat measurements of three randomly assigned patient cases were performed to assess the measurement repeatability for Magnetic Resonance Imaging (MRI), linear array ultrasound (LUS) and phased array ultrasound (PUS). All results are percentages

measurements except for the area of regions 3, 4, 5 and 6 (rostral body, anterior midbody, posterior midbody and isthmus). The correlation strength generally increased with the measurement scale. Stronger correlations to MRI measurements were observed for LUS than for PUS images.

The Bland–Altman plots' 95% confidence intervals were of the order of the size of the measurement values, or larger, for sectional area measurements of MRI compared to LUS and PUS. Confidence intervals were found to be proportionally smaller for the length and total area measurements. Comparable performance was observed for both PUS and LUS compared with MRI data sets. All Bland–Altman plots are available on request.

The lowest intra-observer error, estimated by the COV, for MRI, PUS and LUS was seen for measurements of the long axis length and the total area, while measurements of the area of the rostrum and genu showed the largest COV, approaching 20%. A comparable level of performance could be seen in each of the modalities (Table 2). However, the LUS results showed the lowest COV for the majority of measurements.

DISCUSSION

Our study shows that measurements made from NS are comparable with measurements made from MR images for many structural measurements of the CC and may hence be reliably used to characterise the structure of the CC. Length and total area measurements showed higher correlation strengths, the majority of which were significant at the P value <0.001 level, compared to the sectional areas of the Witelson segmentation. Sectional area measurements showed a wide spread of correlation strengths, half of which were poor ($\rho < 0.4$). Bland–Altman analysis also showed that the tightest 95% confidence intervals, relative to the measurement magnitude, were seen for the length and total area measurements, suggesting that they were more strongly related. The clinical significance of the limits of agreement is unknown.

The higher correlation strengths, smaller confidence intervals and greater repeatability for the length measurements, especially the long axis length, and total area, make these the most suitable candidates for future study, especially considering their potential medical implications.[19,21–24] This improved performance may in part be due to the small size of many of the regional areas which

will have a larger proportional measurement uncertainty. However, correlation strength does not simply improve consistently with increases in area (see Table 2). All regional areas of the CC in this study are defined in relation to the long axis line, which is produced manually.[20] If the placement of this line varies over the three imaging modalities, the measured areas will change. The long axis length measurement is perhaps the most robust measurement explored here; however, even small changes in its angulation may have large implications to the regional areas that result.

The repeatability of any measurement is important to ensure that the conclusions drawn from a set of results can be applied to multiple patients. The results of the intra-observer error testing are encouraging with all pooled COVs <20%, with comparable levels of repeatability observed across all imaging techniques. The long axis length and total area show the greatest level of repeatability. Greater repeatability was seen compared to that reported by Davies et al.,[25] who assessed inter- and intra-observer variability of measurements of the width of the anterior horn, thalamo-occipital distance of the lateral ventricles and widths of the third and fourth ventricles assessed via NS in neonates. While these are neither the same cranial structures nor a matched cohort, they are of similar scales to those measured here. Future research should include a robust assessment of intra- and inter-observer error of measurements obtained from clinical members of staff more accustomed to reviewing neonatal neuroimages.

It had been hypothesised that the LUS data would yield closer correspondence to their MRI counterparts given the improved resolution possible with a less divergent beam and, in this instance, marginally increased central probe frequency. This study appears to add weight to this hypothesis as stronger correlations between length measurements (except for genu width) and total area, the most reliable measurements investigated, were observed when the LUS data was compared to MRI. Further testing may establish whether this apparent improvement is sustained when confounding variables, such as differences in probe frequency, are controlled for.

Throughout this study, all segmentation and measurement of the CC has been performed manually. This is a lengthy and difficult process, which may introduce increased measurement variability and uncertainty. The use of automatic measurement and segmentation techniques, of which recent research demonstrates the viability for some features of the foetal or neonatal brain,[26–28] should be explored in future work: the results of which may make measurement of CC features more feasible in the clinic.

This study was entirely retrospective and hence there was no control of imaging protocols. Future work would benefit from bespoke imaging protocols to ensure that effective imaging of the CC specifically is achieved. This may include altering imaging parameters such as the use of a higher central probe frequency (e.g.12 MHz) for both LUS and PUS imaging, the use of adaptive image processing to enhance boundary detection and spatial compounding to reduce visual clutter that may obscure the edges of the CC for US images. The neonatal brain has a much higher water content than the adult and the myelination of neurons is far from complete, both of which can affect image contrast.[29,30] In order to overcome the T1 and T2 lengthening associated with the increased water content, and hence to recover better contrast in the brain, increases in the TE and TR values used in the MRI studies may also result in improved measurement precision.[29,30] Increased field strength may be beneficial; however, the corresponding safety implications must be considered.

It should be noted that the patient cohort, term neonates with a diagnosis of HIE, are not a cohort representative of the entire neonatal population. However, the increased risk of neurological impairment in this study group makes them a reasonable initial study group.[31]

Comparison of MRI and neurosonogram 1- and 2-dimensional morphological...
M Mills et al.

359

CONCLUSION

The development of the CC is associated with multiple neurological impairments. The ability to identify patients as being at increased risk of specific neurological disorders would allow the care pathway for these patients to be optimised. The use of a bedside, cost-effective, non-invasive and non-ionising imaging modality would be ideal for early diagnostic/prognostic assessment.

This report is novel as data from both linear and phased US arrays is compared with MRI data sets. In this study, we demonstrated that some measurements of the neonatal CC obtained by US images, especially those with a linear array transducer, correlate well with those from MRI. Long axis length and total area appear the best correlated and most precise measurements in this study. These results suggest that some measurements are possibly interchangeable between MRI and US, and hence the biological correlates currently found in studies utilising MRI may be expected from US studies. The feasibility and quality of large-scale research may therefore be increased and improved by focussing on measurements of the CC long axis length and cross-sectional area. The use of imaging protocols designed to optimise the definition and contrast of the CC and employing automatic segmentation and analysis should also be explored.

ADDITIONAL INFORMATION

Competing interests: The authors declare no competing interests.

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

REFERENCES

1. Waxman, S. G. *Clinical Neuroanatomy* 26th edn (McGraw Hill, New York, NY, 2010).
2. Gazzaniga, M. S. Cerebral specialization and interhemispheric communication. Does the corpus callosum enable the human condition? *Brain* **123**, 1293–1326 (2000).
3. Aboitiz, F., López, J. & Montiel, J. Long distance communication in the human brain: timing constraints for inter-hemispheric synchrony and the origin of brain lateralization. *Biol. Res.* **36**, 89–99 (2003).
4. Aboitiz, F. & Montiel, J. One hundred million years of interhemispheric communication: the history of the corpus callosum. *Braz. J. Med. Biol. Res.* **36**, 409–420 (2003).
5. Mihrshahi, R. The corpus callosum as an evolutionary innovation. *J. Exp. Zool. B Mol. Dev. Evol.* **306**, 8–17 (2006).
6. Paul, L. K. Developmental malformation of the corpus callosum: a review of typical callosal development and examples of developmental disorders with callosal involvement. *J. Neurodev. Disord.* **3**, 3–27 (2011).
7. Xia, Z., Hoeft, F., Zhang, L. & Shu, H. Neuroanatomical anomalies of dyslexia: disambiguating the effects of disorder, performance, and maturation. *Neuropsychologia* **81**, 68–78 (2016).
8. Just, M. A., Cherkassky, V. L., Keller, Ta, Kana, R. K. & Minshew, N. J. Functional and anatomical cortical underconnectivity in autism: evidence from an fmri study of an executive function task and corpus callosum morphometry. *Cereb. Cortex* **17**, 951–961 (2007).
9. Arnold, L. E., Hodgkins, P., Kahle, J., Madhoo, M. & Kewley, G. Long-term outcomes of ADHD: academic achievement and performance. *J. Atten. Disord*. https://doi.org/10.1177/1087054714566076 (2015).
10. Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L. & Francis, D. J. The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Read. Res. Q* **40**, 148–182 (2005).
11. Mcgoey, K. E., Eckert, T. L., Dupaul, G. J. & Hyperactivity, A. T. Early intervention for preschool-age children with ADHD: a literature review. *J. Emot. Behav. Disord.* **10**, 14–28 (2002).
12. Wezel-meijler, G. Van, Steggerda, S. J. & Leijser, L. M. Cranial ultrasonography in neonates: role and limitations. *Semin. Perinatol.* **34**, 28–38 (2010).
13. Meijler, G. *Neonatal Cranial Ultrasonography* 2nd edn (Springer, Berlin, 2012).
14. Miller, S. P. et al. Comparing the diagnosis of white matter injury in premature newborns with serial MR imaging and transfontanel ultrasonography findings. *Am. J. Neuroradiol.* **24**, 1661–1669 (2003).
15. Maalouf, E. F. et al. Comparison of findings on cranial ultrasound and magnetic resonance imaging in preterm infants. *Pediatrics* **107**, 719–727 (2001).
16. Rademaker, K. J. et al. Neonatal cranial ultrasound versus MRI and neurodevelopmental outcome at school age in children born preterm. *Arch. Dis. Child. Fetal Neonatal Ed.* **90**, F489–F494 (2005).
17. Leijser, L. M. et al. Structural linear measurements in the newborn brain: accuracy of cranial ultrasound compared to MRI. *Pediatr. Radiol.* **37**, 640–648 (2007).
18. Gilmore, J. H. et al. Infant cerebral ventricle volume: a comparison of 3-D ultrasound and magnetic resonance imaging. *Ultrasound Med. Biol.* **27**, 1143–1146 (2001).
19. Liu, F. et al. Ultrasound measurement of the corpus callosum and neural development of premature infants. *Neural Regen. Res.* **8**, 2432 (2013).
20. Witelson, S. Hand and sex differences in the isthmus and genu of the human corpus callosum: a postmortem morphological study. *Brain* **112**, 799–835 (1989).
21. Thompson, D. K. et al. Accelerated corpus callosum development in prematurity predicts improved outcome. *Hum. Brain Mapp.* **36**, 3733–3748 (2015).
22. Anderson, N. G., Laurent, I., Cook, N., Woodward, L. & Inder, T. E. Growth rate of corpus callosum in very premature infants. *Am. J. Neuroradiol.* **26**, 2685–2690 (2005).
23. Ghassabian, A. et al. Infant brain structures, executive function, and attention deficit/hyperactivity problems at preschool age. A prospective study. *J. Child Psychol. Psychiatry* **54**, 96–104 (2013).
24. Handel, L. L. E. V. A. N. et al. Corpus callosum size in relation to motor performance in 9- to 10-year-old children with neonatal encephalopathy. *Pediatr. Res.* **63**, 103 (2008).
25. Davies, M. W., Swaminathan, M., Chuang, S. L. & Betheras, F. R. Reference ranges for the linear dimensions of the intracranial ventricles in preterm neonates. *Arch. Dis. Child. Fetal Neonatal Ed.* **82**, F218–F223 (2000).
26. Huang, R., Namburete, A. & Noble, A. Learning to segment key clinical anatomical structures in fetal neurosonography informed by a region-based descriptor. *J. Med. Imaging* **5**, 014007 (2018).
27. Beare, R. J. et al. Neonatal brain tissue classification with morphological adaptation and unified segmentation. *Front. Neuroinform.* **10**, 12 (2016).
28. Qiu, W. et al. Automatic segmentation approach to extracting neonatal cerebral ventricles from 3D ultrasound images. *Med. Image Anal.* **35**, 181–191 (2017).
29. Saunders, D. E., Thompson, C., Gunny, R., Jones, R. & Cox, T. Magnetic resonance imaging protocols for paediatric neuroradiology. *Pediatr. Radiol.* **37**, 789–797 (2007).
30. Rutherford, M. A. MRI of the neonatal brain [Internet]. http://www.mrineonatalbrain.com (2002).
31. Jacobs, S. E. et al. Cochrne review: Cooling for newborns with hypoxic ischaemic encephalopathy. *Evid. Based Child Health* **5**, 474–531 (2010).