MRI-based brain volumes of preterm infants at term: a systematic review and meta-analysis

Julia Romberg, Marko Wilke, Christoph Allgaier, Thomas Nägele, Corinna Engel, Christian F Poets, Axel Franz

ABSTRACT
Background MRI allows a detailed assessment of brain structures in preterm infants, outperforming cranial ultrasound. Neonatal MR-based brain volumes of preterm infants could serve as objective, quantitative and reproducible surrogate parameters of early brain development. To date, there are no reference values for preterm infants’ brain volumes at term-equivalent age.

Objective Systematic review of the literature to determine reference ranges for MRI-based brain volumes of very preterm infants at term-equivalent age.

Methods PubMed Database was searched on 6 April 2020 for studies reporting MR-based brain volumes on representative unselected populations of very preterm and/or very low birthweight infants examined at term equivalent age (defined as 37–42 weeks mean postmenstrual age at MRI). Analyses were limited to volumetric parameters reported in >3 studies. Weighted mean volumes and SD were both calculated and simulated for each parameter.

Results An initial 367 publications were identified. Following application of exclusion criteria, 13 studies from eight countries were included for analysis, yielding four parameters. Weighted mean total brain volume was 379 mL (SD 72 mL; based on n=756). Cerebellar volume was 21 mL (6 mL; n=791), cortical grey matter volume 140 mL (47 mL; n=572) and weighted mean volume of unmyelinated white matter was 195 mL (38 mL; n=499).

Conclusion This meta-analysis reports pooled data on several brain and cerebellar volumes which can serve as reference for future studies assessing MR-based volumetric parameters as a surrogate outcome for neurodevelopment and for the interpretation of individual or cohort MRI-based volumetric findings.

INTRODUCTION
Preterm birth is a major contributor to the global burden of disease.1 Neurodevelopmental outcome after preterm birth ultimately determines quality of life; however, the determinants of this outcome are still subject to research.2–5

Impaired neurodevelopmental outcome may be present without overt intracerebral lesions detectable by neonatal cranial ultrasound.6 In order to identify valid biomarkers that facilitate the prediction of neurocognitive long-term outcome and, at the same time, establish early surrogate endpoints for clinical studies aimed at improving perinatal care, various parameters have been studied on cerebral MRI in preterm infants in the last decades. Of these, advanced methods yielding quantitative parameters (such as volumetric MRI, diffusion MRI, 1H-MR spectroscopy and resting-state functional connectivity) seem to be most promising for providing objective, predictive and sensitive values.7

This systematic review focuses on volumetric MRI, aiming at identifying reference ranges for cerebral and cerebellar volumes in unselected very preterm/very low birthweight (VLBW) infants at term-equivalent age. There is evidence suggesting that reduced brain tissue volumes in preterm infants are associated with impaired neurocognitive outcomes.8–13 Reduced cerebral and cerebellar tissue volumes may result from loss of immature oligodendrocytes which are vulnerable to inflammation and fluctuations in haemodynamics and oxygenation eventually resulting in altered cerebral development and impaired function.8–14–17

METHODS
To identify all published studies reporting brain volume measurements of very preterm or VLBW infants, PubMed was searched on 6 April 2020 using the following systematic search strategy: ‘(infant, premature’ (MeSH) OR ‘infant, low birth weight’ (MeSH)) AND ‘magnetic resonance imaging’ (MeSH) AND (volum*(All Fields)). The term ‘brain’ was intentionally not included in the search strategy to avoid omitting studies focusing on, for example, cerebellar/hippocampal volume...
measurements. The process of reviewing and assessing studies was based on the guidelines established by Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).

Articles were selected for data extraction if they fulfilled all of the following criteria:

- Reporting volumetric data of the brain or its anatomical structures acquired by structural MRI.
- Examining very preterm infants (ie, born at a gestational age <32 weeks) or very low birth weight (VLBW) infants (ie, with birth weight <1500 g).
- Reporting MRI data acquired at term-equivalent age (ie, if the mean postmenstrual age (PMA) at scan was between 37 and 42 weeks).
- Including representative populations (ie, studies reporting preselected subpopulations were excluded).
- Published in English.

- Accessible (either online, via interlibrary loan, or by directly contacting the authors).

Further,

- If several manuscripts reported data from the same population, only the study including the largest number of infants was included in this review. If there was only a small overlap between patient populations (<25%), both studies were included.
- Only brain imaging outcome parameters reported in >3 studies were included in the meta-analysis.

First, title and abstract of all articles identified with the search strategy were reviewed manually (by JR and CA) with regard to the above-mentioned inclusion and exclusion criteria, and studies not satisfying these criteria were excluded. In potentially eligible studies, the full text of each article was reviewed. The bibliography of each eligible full-text article was hand-searched for additional studies. If studies reported concurrent populations of term-born infants, brain volume data were extracted and reported, too.

The following quality criteria were assessed and reported for all eligible studies: Were the infants representative of the underlying population? Was the reliability of image processing methods tested?

Cohorts of included studies were assessed regarding the following characteristics: criteria for inclusion and exclusion in the MRI study, site and time period of recruitment, mean birth weight, mean gestational age at birth, proportion of small for gestational age (SGA) infants, use of antenatal and postnatal steroids, proportion of infants with bronchopulmonary dysplasia (BPD), persistent ductus arteriosus (PDA) and intraventricular haemorrhage (IVH), postmenstrual age (PMA) at MRI scan and type of MRI scanner.

Statistical analyses

All volumes are reported in millilitres. Gestational age at birth and PMA at MR scan are reported in weeks. In case data had been reported by subgroups, the weighted mean for the overall study population was calculated based on all subgroups.

Meta-analysis of volumetric data was performed in two ways:

(A) Weighted mean values and weighted SD were calculated, the latter on the basis of total variance using SAS statistical software version 9.4:

\[
\text{Total variance: } s^2_{\text{total}} = \frac{1}{n} \sum_{i=1}^{M} w_i \times s^2_i + \frac{1}{n} \sum_{i=1}^{N} w_i \times (x_i - x^*)^2
\]

(B) Data from all studies for a given parameter were simulated using the Matlab random number generator (The MathWorks, Natick, Massachusetts, USA): Based on the mean (SD) of a particular study, a total of n populations with normally distributed 100 000 data points each were generated, with n corresponding to the number of subjects included in the original studies. These synthetic populations of all studies were combined by means of histogram analyses to yield a final population. To this new population (representing all contributing populations in accordance with their original size), we fitted a Gaussian distribution from which a final population mean and SD, as well as a goodness-of-fit parameter was reported.

Linear regression for brain volumes in relation to mean age at scan was calculated and visualised using Python V.3.7 (https://www.python.org/).

Registration

This review was not registered. A review protocol was prepared, but not published.
Table 1 Demographic variables of study populations

<table>
<thead>
<tr>
<th>Reference</th>
<th>N</th>
<th>Inclusion criteria</th>
<th>Site of recruitment</th>
<th>Study period</th>
<th>Birth weight (g) Mean (SD)</th>
<th>Gestational age (weeks) Mean (SD)</th>
<th>SDS (SDS = (Mean(SD)/10)/% IUGR*</th>
<th>% Antenatal Corticosteroids</th>
<th>% Postnatal Corticosteroids</th>
<th>% BPD</th>
<th>% PDA</th>
<th>% any IVH</th>
<th>PMR at scan (weeks) Mean (SD)</th>
<th>MR scanner</th>
<th>Segmentation method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blok et al., Plas ’t Harde 2014</td>
<td>36</td>
<td>GA &lt;32 weeks</td>
<td>Utrecht (NL)</td>
<td>2007–2009</td>
<td>1027 (47)</td>
<td>27.2 (1.7)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>42%</td>
<td>28%</td>
<td>°IV: 38.9%, °I: 14%, °II: 5%</td>
<td>41 (0.8)</td>
<td>1.5T Philips Achieva</td>
<td>Semi-automatic</td>
</tr>
<tr>
<td>Elbid et al. J Pediatr, 2010</td>
<td>20</td>
<td>bw&lt;1500 g and GA &lt;32 weeks</td>
<td>Turku (FI)</td>
<td>2001–2006</td>
<td>1121 (319)</td>
<td>29.0 (2.7)</td>
<td>−1.43 (1.61)/38% (SGA)</td>
<td>N/A</td>
<td>14%</td>
<td>14%</td>
<td>12.5%, missing information for 4</td>
<td>°IV: 23.7%, °I: 11%</td>
<td>Manual Philips Gyroscan Intera</td>
<td>Manual</td>
<td></td>
</tr>
<tr>
<td>Hakan et al., J Clin Endocrin Metab, 2011</td>
<td>51</td>
<td>GA &lt;31 weeks</td>
<td>Lund (SE)</td>
<td>2005–2007</td>
<td>1121 (319)</td>
<td>29.0 (1.9)</td>
<td>N/A</td>
<td>N/A</td>
<td>26% hydrocortisone</td>
<td>N/A</td>
<td>N/A</td>
<td>42%</td>
<td>41.4 (0.8)</td>
<td>Manual Philips Intera</td>
<td>Manual</td>
</tr>
<tr>
<td>Ekblad et al., J Pediatr, 2010</td>
<td>30</td>
<td>bw&lt;1500 g and GA &lt;32 weeks</td>
<td>Turku (FI)</td>
<td>2001–2006</td>
<td>1121 (319)</td>
<td>29.0 (1.9)</td>
<td>−1.43 (1.61)/38% (SGA)</td>
<td>N/A</td>
<td>14%</td>
<td>14%</td>
<td>12.5% (missing information for 4)</td>
<td>°IV: 23.7%, °I: 11%</td>
<td>Manual Siemens Magnetom Allegra</td>
<td>Manual</td>
<td></td>
</tr>
<tr>
<td>Hansen-Pupp et al., J Clin Endocrin Metab, 2011</td>
<td>22</td>
<td>GA &lt;31 weeks</td>
<td>Lund (SE)</td>
<td>2005–2007</td>
<td>888 (281)</td>
<td>26.4 (1.9)</td>
<td>N/A</td>
<td>24%</td>
<td>26% hydrocortisone</td>
<td>N/A</td>
<td>N/A</td>
<td>24%</td>
<td>°I–IV: 38.9%</td>
<td>Manual Siemens Magnetom Allegra</td>
<td>Manual</td>
</tr>
<tr>
<td>Inder et al., J Pediatr, 2005</td>
<td>8</td>
<td>119 (of 129)**</td>
<td>Auckland (NZ), n=100: Christchurch (NZ), n=29: Melbourne (AU)</td>
<td>1998–2000</td>
<td>1040 (272)</td>
<td>27.7 (2.2)</td>
<td>22%</td>
<td>8%</td>
<td>21%</td>
<td>2%</td>
<td>°I–IV: 17%</td>
<td>40.2 (0.3)</td>
<td>1.5T Philips Achieva/Siemens Avanto</td>
<td>Manual Siemens Magnetom Allegra</td>
<td>Semi-automatic</td>
</tr>
<tr>
<td>Kamino et al., Pediatr Res, 2018</td>
<td>28</td>
<td>44 (of 60)††</td>
<td>San Francisco (USA), n=33: University of California San Francisco (UCSF), n=27: University of British Columbia (UBC)</td>
<td>2010–2011</td>
<td>440 (60)††</td>
<td>28.1 (1.8)</td>
<td>28.5 (2)</td>
<td>100%</td>
<td>28.1%</td>
<td>48%</td>
<td>UCSF: 12%, UBC: 37%</td>
<td>°I–II: 12%, °III–IV: 37%</td>
<td>48%</td>
<td>37.7 (5)</td>
<td>1.5T Siemens Magnetom Allegra</td>
</tr>
<tr>
<td>Moekle et al., Plas ’t Harde 2014</td>
<td>85</td>
<td>GA &lt;28 weeks</td>
<td>Utrecht (NL)</td>
<td>2008–2013</td>
<td>889 (296)</td>
<td>26.0 (1.9)</td>
<td>−1.0 (1.2)</td>
<td>N/A</td>
<td>24%</td>
<td>2%</td>
<td>°I–II: 19%, °III–IV: 4%</td>
<td>°I–IV: 38%, °I: 11%</td>
<td>40.1 (0.6)</td>
<td>1.5T Siemens Magnetom Allegra</td>
<td>Semi-automatic</td>
</tr>
<tr>
<td>Parkh et al., Plas ’t Harde 2014</td>
<td>122</td>
<td>bw&lt;1000 g and GA &lt;32 weeks</td>
<td>Houston (USA)</td>
<td>2005–2007</td>
<td>771 (114)</td>
<td>26.7 (1.9)</td>
<td>25% (SGA)</td>
<td>70%</td>
<td>10%</td>
<td>3%</td>
<td>°I–IV: 38%</td>
<td>38.5 (2.2)</td>
<td>1.5T Philips Achieva/Siemens Avanto</td>
<td>Semi-automatic</td>
<td></td>
</tr>
<tr>
<td>Steinman et al., J Pediatr, 2015</td>
<td>192</td>
<td>bw&lt;1500 g and GA &lt;30 weeks</td>
<td>Melbourne (AU)</td>
<td>2001–2003</td>
<td>956 (223)</td>
<td>27.5 (3.3)</td>
<td>9%</td>
<td>8%</td>
<td>9%</td>
<td>N/A</td>
<td>°I–IV: 38%, °I: 11%</td>
<td>40.1 (0.6)</td>
<td>1.5T Siemens Magnetom Allegra</td>
<td>Semi-automatic</td>
<td></td>
</tr>
<tr>
<td>Silve et al., J Chir Neur, 2019</td>
<td>105</td>
<td>GA &lt;32 weeks</td>
<td>Toronto (CA)</td>
<td>2008–2010</td>
<td>288 (38)</td>
<td>6% (postnatal steroids for BPD)</td>
<td>N/A</td>
<td>52%</td>
<td>3%</td>
<td>PDA therapy</td>
<td>°I–IV: 38%, °I: 11%</td>
<td>40.1 (0.6)</td>
<td>1.5T Siemens Magnetom Allegra</td>
<td>Semi-automatic</td>
<td></td>
</tr>
<tr>
<td>Tam et al., SCI Ramb Med, 2011</td>
<td>68</td>
<td>GA&lt;33 weeks</td>
<td>San Francisco (USA)</td>
<td>2006–2009</td>
<td>27.5 (2.1)</td>
<td>N/A</td>
<td>90%</td>
<td>22%</td>
<td>51%</td>
<td>50%</td>
<td>°I–IV: 12%, °I: 11%</td>
<td>°I: 19%, °II: 21%</td>
<td>43 (0.3)</td>
<td>1.5T Philips Achieva</td>
<td>Semi-automatic</td>
</tr>
<tr>
<td>Thompson et al., Brain, 2017</td>
<td>202</td>
<td>bw&lt;1500 g and GA&lt;30</td>
<td>Melbourne (AU)</td>
<td>2001–2004</td>
<td>965 (219)</td>
<td>27.6 (1.1)</td>
<td>11%</td>
<td>8%</td>
<td>36% (Indomethacin administered)</td>
<td>°I–IV: 38%, °I: 11%</td>
<td>40.1 (0.6)</td>
<td>1.5T Siemens Magnetom Allegra</td>
<td>Semi-automatic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vasu et al., BMJ Open, 2014</td>
<td>19</td>
<td>19 (of 22)**</td>
<td>London (UK)</td>
<td>2007–2008</td>
<td>1260 (410)</td>
<td>28.9 (2.8)</td>
<td>−0.13 (0.76)</td>
<td>N/A</td>
<td>N/A</td>
<td>25%</td>
<td>°I–IV: 38%, °I: 11%</td>
<td>40.1 (0.6)</td>
<td>1.5T Philips Achieva</td>
<td>Automatic</td>
<td></td>
</tr>
</tbody>
</table>

N/A: Data not available in the published manuscript.

*Brain volume was not available in all infants.

†Proportion with SDS for birth weight <-2 if marked as 'SGA', then proportion with birth weight <-10th percentile.


¶From the original data reported by Tam et al., the infants meeting the inclusion criteria of this review were selected and analysed.

**Only the percentage of infants with hydrocortisone treatment was reported.

††Patient characteristics were reported from the original population.

bw, birth weight; GA, gestational age at birth; IVH, intraventricular haemorrhage; SGA, small for gestational age; UBC, University of British Columbia; UCI, University of California in San Francisco; WMD, white matter disease.
RESULTS
Using our initial search criteria, a total of 367 studies, published between 1992 and 2019, were retrieved. After review for inclusion and exclusion criteria, 13 were included (see figure 1). Hand-searching the bibliographies yielded two additional studies meeting the inclusion criteria.18 19 However, both could not be included because volume data were reported as median (IQR) and least-squares means, respectively. This left 13 articles for inclusion in the meta-analysis.

We identified potential overlap in populations between studies.20–23 In these cases, for each volumetric parameter, only the study with the largest population was included for the respective analysis. A small overlap in the population of two studies8 25 also resulted in 29 infants being included twice.

Characteristics of the included study populations are listed in table 1. Studies indicated the following reasons for exclusion of subjects: Major congenital anomalies (stated in 9 out of 13 studies) and low-quality MR images (stated in 5 out of 13 studies). All studies excluded infants with missing parental consent. Three studies additionally excluded infants with congenital infection (no further definition given).

Only one study tested representativeness of its study population, and there were no differences between included and excluded subjects regarding gender, multiple birth, gestational age, IVH and BPD.25 Some studies indicated recruitment rates in relation to eligible subjects, which varied considerably (90%/80%, 21% 73%, 21% 100%, 29% 66%, 24% 94%). Most studies (10/13) were single-centre cohorts from level III academic centres in highly developed countries.

All studies reported MRI postprocessing and segmentation methods. Except for one study,31 all studies specified reliability testing for their technique. Studies employed manual, semiautomated and in three studies,20 21 29 automated segmentation techniques (see table 1).

Brain volume measurements
The following volumetric parameters were reported in >3 studies: total brain volume (TBV; n=756), cerebellar volume (CV; n=791), volume of cortical grey matter (CGM; n=572) and volume of unmyelinated white matter (UMWM; n=499).

Brain volumes are presented in table 2. Figure 2 displays the relation between volume measurements and mean PMA at scan. Simulated populations and fitted Gaussian distribution are depicted in figure 3.

Four studies21 25 29 31 also reported MRI-based brain volumetry in contemporary samples of term-born infants. Results are summarised in online supplemental table S1 and indicate that CGM and UMWM volumes might be reduced by about 25% in preterm compared with term infants.

Based on the simulated population data, an online supplemental MS-excel spread sheet is provided converting MRI brain volumes of preterm infants into z-scores.

<p>| Table 2 | Brain volume measurements (and weighted means) in preterm infants at term-equivalent age |
|-----------------|---------------------------------|-----------------|-------------------|-------------------|------------------|-------------------|-----------------|</p>
<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>PMA at scan</th>
<th>Total brain volume</th>
<th>Total cerebellar volume</th>
<th>Cortical grey matter volume</th>
<th>Unmyelinated white matter volume</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(weeks) Mean (SD)</td>
<td>(mL) Mean (SD)</td>
<td>(mL) Mean (SD)</td>
<td>(mL) Mean (SD)</td>
<td>(mL) Mean (SD)</td>
</tr>
<tr>
<td>Blok et al, 201420</td>
<td>36</td>
<td>41.4 (0.8)</td>
<td>391 (34)</td>
<td>N/A</td>
<td>167 (23)* (overlap with Moeskops et al31)</td>
<td>161 (21)* (overlap with Moeskops et al31)</td>
</tr>
<tr>
<td>Ekkblad et al, 201020</td>
<td>209</td>
<td>t</td>
<td>397 (51.2)</td>
<td>24.2 (5.1)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Hansen-Pupp et al, 201122</td>
<td>51</td>
<td>40.1 (0.6)</td>
<td>396 (46.3)* (n=46, overlap with Hansen-Pupp et al31)</td>
<td>21.3 (2.7)</td>
<td>N/A</td>
<td>191 (25)* (n=46, overlap with Hansen-Pupp et al31)</td>
</tr>
<tr>
<td>Hansen-Pupp et al, 201322</td>
<td>49</td>
<td>40.1 (0.6)</td>
<td>393 (50.9)</td>
<td>21.1 (3.5)* (overlap with Hansen-Pupp et al31)</td>
<td>N/A</td>
<td>190 (25)</td>
</tr>
<tr>
<td>Inder et al, 20058</td>
<td>119</td>
<td>40.2 (0.3)</td>
<td>406 (57)</td>
<td>N/A</td>
<td>178 (41)</td>
<td>202 (41)</td>
</tr>
<tr>
<td>Kamino et al, 201828</td>
<td>44</td>
<td>37.7 (3.0)</td>
<td>N/A</td>
<td>16.7 (5.3)</td>
<td>107.6 (40.2)</td>
<td>151 (20.1)</td>
</tr>
<tr>
<td>Moeskops et al, 201521</td>
<td>85</td>
<td>41.1 (0.5)</td>
<td>N/A</td>
<td>107 (13)</td>
<td>171 (19)</td>
<td></td>
</tr>
<tr>
<td>Parikh et al, 201323</td>
<td>122</td>
<td>38.5 (2.2)</td>
<td>270 (41.5)</td>
<td>16.1 (4.2)</td>
<td>105 (19.1)</td>
<td>N/A</td>
</tr>
<tr>
<td>Steinhorn et al, 201524</td>
<td>192</td>
<td>t</td>
<td>395 (63.2)* (overlap with Thompson et al25)</td>
<td>21.4 (4.5)</td>
<td>159 (41)* (overlap with Thompson et al25)</td>
<td>N/A</td>
</tr>
<tr>
<td>Stiver et al, 201526</td>
<td>105</td>
<td>41.9 (2.0)</td>
<td>N/A</td>
<td>23.9 (5)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Tam et al, 201127</td>
<td>68</td>
<td>39.5 (1.3)</td>
<td>N/A</td>
<td>20.4 (5)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Thompson et al, 200728</td>
<td>202</td>
<td>40.1 (1.7)</td>
<td>395 (64)</td>
<td>N/A</td>
<td>159 (41)</td>
<td>212 (32)</td>
</tr>
<tr>
<td>Vasu et al, 201429</td>
<td>19</td>
<td>40.2 (3.1)</td>
<td>462 (12.8)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Aggregate data all studies Weighted mean (Total variance)</td>
<td>379 (72); n=756</td>
<td>21 (6); n=791</td>
<td>140 (47); n=572</td>
<td>195 (38); n=499</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggregate data all studies Mean (SD) by simulation</td>
<td>391 (67)†</td>
<td>19 (5)§</td>
<td>135 (68)§</td>
<td>192 (38)§</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggregate data excluding reference29 Weighted mean (Total variance)</td>
<td>399 (57) n=634</td>
<td>22.3 (5.3) n=669</td>
<td>149.2 (47.8) n=450</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
DISCUSSION

For this systematic review, we identified 13 studies reporting MRI-based brain volume measurements in very preterm infants at term-equivalent age. For TBV, CV, CGM and UMWM, sufficient data were available for this meta-analysis.

All studies included infants with neonatal complications such as BPD, PDA and IVH. Since exclusion criteria of the studies were limited to missing parental consent, ‘congenital anomalies’ and low-quality MRI images and recruitment rates were ≥67% (reported in six studies), we can assume these study populations to be representative of very preterm infants. However, most studies included single-centre cohorts from academic level III centres in highly developed countries, limiting generalisability to some extent.

In one study, inclusion was restricted to infants ≤1000 g birth weight; consequently, mean birth weight was lower than in the other study populations. Mean TBV, CV and CGM volume all were considerably lower in this cohort compared with the other preterm populations. This discrepancy can partly be attributed to the fact that the timepoint of MRI measurements in this study (mean PMA at scan: 38.5±2.2 weeks) was rather early compared with the other studies. Another explanation might be that ELBW infants (i.e. birthweight < 1000g) are more prone to impaired brain growth than VLBW infants; these hypotheses should be tested in future studies or individual patient data meta-analyses. A number of studies (including that of Parikh et al[29]) found smaller brain volumes in IUGR/SGA infants. 8 22 25 29 32 33 Notably, Parikh et al also found brain volumes of contemporary term born infants to be markedly smaller compared with those reported by other studies. Thus, a systematic measurement bias could also have contributed to the observed differences. For reference purposes, we also provide the weighted mean and SD values excluding Parikh’s study for TBV, CV and CGM volume.

Figure 2 visualises the variance of measured brain volumes among the studies, highlighting the need of large study groups and pooled data. Additionally, linear regression was calculated for brain volumes by PMA at scan showing the best fit for a linear relation for CV. Measurements of CV could be more reliable due to clear anatomical boundaries, resulting in the lower variability. Furthermore, studies reporting CV were well dispersed for PMA at scan. According to Gui et al,34 the cerebellum shows one of the highest growth rates around term, which could explain why a growth trend can be illustrated here. For the other parameters, variance in this restricted time frame is too high to deduce growth trends from the data.

This systematic review and meta-analysis has several limitations: Only one database (PubMed) was searched and only articles in English were included. However, only a very small number
considerable variability in image acquisition and processing was recognised. For brain volume extraction, tissue segmentation into grey and white matter and cerebrospinal fluid (CSF) is the crucial step. Of the included studies, five used a manual, five a semiautomatic and three an automatic approach (cf. table 1). All but one study reported internal validation of their segmentation process. While a congruent approach would of course have been preferable, this (incidentally rather well-balanced) composition makes a substantial bias resulting from our data processing approach unlikely.

Despite heterogeneity concerning study procedures and reporting practices, relatively congruent brain volumes were reported, apart from the ELBW study. In total, a substantial number of infants could be included in the meta-analysis. It is also reassuring that both methods used to combine the values across populations, despite taking very different approaches, also yielded very similar values (cf. table 2). Thus, the resulting mean values and SDs presented in this review can serve as reference data for future volumetric MRI studies among preterm infants and can be consulted when interpreting individual or cohort brain volume measurements in VLBW/very preterm infants. As online supplemental data, we provide a calculator embedded within an Excel worksheet for easy comparison of individual versus these group values.

Ultimately, brain volumes of term born infants should serve as reference for regular, intrauterine brain growth. However, despite all efforts, there is still a gap between cohorts of term born and very preterm infants. In 4 of 13 studies included in this review, concurrent cohorts of term infants were scanned around term-equivalent age, and significantly smaller total brain tissue volumes were demonstrated for preterm infants in all four studies. According to Thompson et al and Inder et al, this reduction is driven by smaller CGM and DGM volumes. Furthermore, all four studies revealed significantly increased total CSF volumes in preterm infants, likely reflecting loss or inadequate growth of cerebral tissue (ie, grey or white matter).

Although brain growth of term born infants should be the ultimate goal, brain volume data of unselected very preterms are required when planning interventional studies aiming at optimising brain growth on a realistic basis.

As mentioned above, several studies found an association between birth weight and brain volumes at term. Additionally, male sex and higher gestational age at birth were found to be associated with larger brain volumes. While it would have been interesting to separate out brain volumes with regard to these factors, the respective data were not available in sufficient detail to allow for their inclusion in multivariate analyses. Future studies on MRI brain volumes in VLBW/very preterm infants should aim at further identifying relevant factors that influence early brain volume growth after preterm birth and, if possible, at quantifying their impact. In this context it would be of special interest to address those factors that are not determined (such as sex) but can be modified by optimised neonatal intensive care (such as nutrition, oxygenation, etc). Further studies are also needed that investigate the long-term predictive value of individual MR-based brain volume measurements at term with respect to neurocognitive outcome in childhood and beyond.

In conclusion, in this meta-analysis of 13 studies including data from more than 900 very preterm infants, with the exception of one study, rather similar results for TBV, CV and UMWM volumes were found at term-equivalent age. These weighted mean volumes may serve as reference for individual patient data and future studies.
Correction notice This article has been corrected since it was first published. The open access licence has been updated to CC BY.

Contributors AF, MW and JR conceived the idea of this systematic review. AF and JR developed the research strategy and criteria of inclusion and exclusion. JR and CA performed the research and selection of studies. AF, MW and JR devised the concept of data extraction, calculation and presentation. MW, CA, CE and JR performed the calculations. JR wrote the manuscript with support of MW, CPF, AF, CA and TN, who all made important intellectual contributions and approved the final version of the manuscript. MW and AF planned and designed the supplemental material. JR acts as guarantor for the study.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study does not involve human participants.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information. The authors confirm that the data supporting the findings of this study are available within the article and/or its supplementary materials.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any errors or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 International license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.

ORCID ids
Julia Romberg http://orcid.org/0000-0003-3994-8052
Christian F Poets http://orcid.org/0000-0002-1072-0066
Axel Frenz http://orcid.org/0000-0001-5066-0857

REFERENCES