Check for updates

OPEN ACCESS

EDITED BY Juri Taborri, University of Tuscia, Italy

REVIEWED BY Simone Pasinetti, University of Brescia, Italy Chenguang Zhang, Sanofi Genzyme, United States

*CORRESPONDENCE Ronald van Gils, ⊠ r.vangils@erasmusmc.nl

RECEIVED 30 August 2024 ACCEPTED 28 October 2024 PUBLISHED 13 November 2024

CITATION

van Gils R, Helder O, Kornelisse R, Reiss I and Dankelman J (2024) 3D scanner measuring preterm infants' head circumference and cranial volume: validation in a simulated care setting. *Front. Med. Eng.* 2:1463793. doi: 10.3389/fmede.2024.1463793

COPYRIGHT

© 2024 van Gils, Helder, Kornelisse, Reiss and Dankelman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

3D scanner measuring preterm infants' head circumference and cranial volume: validation in a simulated care setting

Ronald van Gils^{1,2,3,4}*, Onno Helder^{2,3}, René Kornelisse¹, Irwin Reiss¹ and Jenny Dankelman⁵

¹Department of Neonatal and Pediatric Intensive Care, Division of Neonatology, Erasmus MC University Medical Center, Rotterdam, Netherlands, ²Department of Create4Care, Erasmus MC University Medical Center, Rotterdam, Netherlands, ³Research Centre Innovations in Care, Rotterdam University of Applied Sciences, Rotterdam, Netherlands, ⁴Institute of Engineering and Applied Science, Rotterdam University of Applied Sciences, Rotterdam, Netherlands, ⁵Department of Biomechanical Engineering, Faculty of Mechanical Engineering, Delft University of Technology, Delft, Netherlands

Introduction: Weekly head circumference (HC) measurements using a measuring tape is the current standard for longitudinal brain growth monitoring of preterm infants. The MONITOR3D (M3D) 3D scanner has been developed to measure both HC and cranial volume (CrV) of preterm infants within incubators. The M3D's usability, accuracy and precision were validated in a simulated setting in a neonatal intensive care unit (NICU).

Materials and methods: During a simulated routine care moment, NICU nurses conducted M3D scans of a preterm doll simulating an extreme low birthweight preterm (ELBW; BW < 1,000 g) infant, followed by manual HC measurements using a measuring tape. Usability was quantified by percentage of successful HC and CrV measurements from scans. HC and CrV were calculated by marking anatomical landmarks on the 3D image. Measurements were compared to the real, ground truth (GT) values of the doll's head, defined by an accurate medical scanner. Measurement accuracy was assessed using mean or median absolute measurement error (ME), and precision by the spread of ME, represented by the 95% interval of the ME range. ME intervals were compared with preterm weekly growth increases to assess clinical usability.

Results: Regarding usability, 56 M3D scan sessions resulted in 25 successful (44.6%) HC and CrV measurements, with incomplete 3D data being the primary cause of unsuccessful scans. Accuracy of the measuring tape for HC was 0.2 cm (proportional 0.9% of GT), and precision was 1.6 cm (6.3%). M3D's accuracy of HC was 0.4 cm (1.5%), and precision was 0.7 cm (2.9%). For CrV, M3D's accuracy was 8.0 mL (3.8%) and precision 22.6 mL (10.8%).

Conclusion: The M3D scanner is suitable for measuring HC and CrV in ELBW infants. However, current scan success rate is too low for practical usability. The M3D's accuracy and precision are clinically sufficient, while the precision of the current measuring tape method is inadequate for preterm infants. This makes the M3D a promising alternative for HC, offering less disturbance to the infant. In the

future, the M3D technique could facilitate the creation of CrV growth reference charts for ELBW infants, enhancing the accuracy of clinical growth monitoring for preterm infants.

KEYWORDS

neonatal intensive care unit (NICU), preterm infants, extremely low birth weight (ELBW), growth monitoring, head circumference (HC), cranial volume (CrV), 3D scanning and measuring, neonatal stress reduction

Introduction

In current clinical practice, longitudinal brain growth monitoring of preterm infants in neonatal intensive care units (NICUs) is typically performed by weekly measuring head circumference (HC) using a measuring tape wrapped around the head (Villar et al., 2015; Villar et al., 2013). However, this method has several limitations, including inducing stress and providing a poor estimation of brain volume (Vermeulen et al., 2021; Burkhardt et al., 2019; Santander et al., 2020; Martini et al., 2018; Meyer-Marcotty et al., 2018; Ifflaender et al., 2013). Stress caused by painful or uncomfortable procedures is associated with suboptimal brain development (Ranger and Grunau, 2014; Brummelte et al., 2012) and may lead to negative health outcomes (Ranger et al., 2013; Grunau, 2013; Doesburg et al., 2013; Grunau et al., 2009; Ranger et al., 2015). The widespread recognition of the adverse effects of neonatal stress had led to the adoption of individualized care programs focused on stress prevention/reduction, which are now standard of care in NICUs (Moody et al., 2017; Als et al., 2012; VandenBerg, 2007).

Multiple studies have shown that 3D cranial volume (CrV) is a more accurate predictor for brain development compared to HC (Vermeulen et al., 2021; Burkhardt et al., 2019; Santander et al., 2020; Martini et al., 2018; Meyer-Marcotty et al., 2018; Ifflaender et al., 2013). However, there is no CrV data available that can serve as growth reference data, comparable to the internationally acknowledged growth reference charts for preterm HC (Young et al., 2020; RCPCH, 2019; Villar et al., 2016). The lack of preterm CrV data is mainly due to the many technical and practical challenges of stress-free measurement in ventilated infants within incubators. Few studies have aimed to establish longitudinal CrV data of preterm infants. Vermeulen et al. presented the first CrV growth reference charts for preterm infants of 34–42 weeks gestational age (GA) and recommended developing measuring techniques that would enable gathering CrV data for younger enfants (Vermeulen et al., 2021). Dieks et al. presented longitudinal HC and CrV data for infants of 28–40 weeks GA (Dieks et al., 2022).

An accurate, stress-free technique is needed for CrV measurements of preterm infants within incubators during routine care. Such a technique could establish CrV reference data for even extremely low birth weight (ELBW; BW < 1,000 g) infants. To achieve this, The Dutch company Vectory3 (Vectory3, The Hague, Netherlands) developed the MONITOR3D (M3D) scanner (Figure 1) in corporation with the Erasmus University



FIGURE 1

The MONITOR3D scanner, devised as a portable device, is placed on the incubator's transparent cover and operated via a connected laptop on a laptop stand. The three depth cameras are positioned closely around the transparent cover. The doll's hat is removed before scanning. When the 'Take Capture' button is activated on the laptop touchscreen, three (overlapping) 3D images are captured simultaneously, one from each camera, with a capture time of 30 msec. No other handling with the M3D scanner/laptop is required other than activating the capture button on the laptop. After the capture (in a fraction of sec), nurses can instantly continue with their routine care handling. The only extra handling, caused by the M3D, is that nurses should enable a clear view of the top part of the head, so not obstructed by any hat, tube, blanket, etc.

Medical Center (Zevenbergen, 2020). This study aims to validate the usability, accuracy and precision of the novel M3D scanner in a simulated NICU-care setting.

Materials and methods

Design and setting

We conducted a validation study in a simulated setting in June 2022 at the NICU of the Erasmus MC–Sophia Children's Hospital. This NICU comprises four open baby sub-units, each with eight beds, and admits approximately 600 neonates annually. During the study period, the clinical staff included 17 neonatologists, nine residents, 91 nurses, 16 nurse trainees, 22 nursing assistants, and nine nurse practitioners.

Before the study, the NICU staff was informed through a clinical lecture and an announcement in the NICU newsletter. The validation took place at an NICU incubator bedspace using a prepared preterm doll lying in an incubator (Figure 1), simulating a ELBW infant weighing approximately 690 g and measuring 35 cm in length (SOMSO Premature Baby, Female, SOMSO MODELLE GmbH, Coburg, Germany). The doll was equipped with a ventilation tube and a feeding tube both secured to the nose and forehead with commonly used tape, and was positioned in a supportive 'snuggle' with a rolled-up towel around the legs (Figure 2).

Research ethic committee approval

The Clinical Technological Research-Human Research Ethics Committee (KTO-HREC) of Delft University of Technology approved the study (Application number KTO-HREC21005). All collected data were stored anonymously.

The MONITOR3D scanner

The M3D scanner used in this study was in the prototyping stage and not commercially available (Zevenbergen, 2020). The scanner consists of three (eye-safe) 3D cameras (Intel RealSense, Santa Clara, California, USA) assembled in fixed positions within a rigid frame that fits precisely over the transparent cover of a Caleo incubator (Dräger, Lübeck, Germany), positioning the three 3D cameras around the head. (Figure 1).

Obviously, the cameras cannot capture the part of the head touching the matrass. To capture all sides of the head, three separate scan sessions are needed, each with the head in a different position, facing up, right and left. The three scans result in a total of nine overlapping 3D images, covering all areas of the head.

To prevent stress from the head position changes required for the scans, the three sessions are to be synchronized with the protocolized head position changes, which are part of standard NICU standard care. These position changes occur every three to 4 hours, aiming to prevent skull deformations (Kubik et al., 2024; De Koning, 2020). Needed three capture should be done on the same day, during two or maximum three successive routine care moments. If intervals between captures are too long, the head could be significantly modified by growth, creating bias in the measurement. The head position change is integrated with routine care procedures, such as diaper change. By aligning the M3D scan sessions with the routine care moments, the first scan is conducted with the head in the initial position at the start of the routine care. Subsequently, routine care procedures are performed, concluding with the infant in a different head position for the second M3D scan. After three to 4 hours, the infant's head is repositioned once, providing an opportunity for the third scan.

Data collection

Data collection involved three steps:

Scan sessions

Pairs of NICU nurses conducted scan sessions using the M3D scanner with a preterm doll (Figure 2) placed in an incubator (Figure 3), specially prepared to simulate a NICU routine care moment. Researchers first demonstrated how to operate the M3D and clarified expectations. Each complete scan session entailed three separate 3D-scans, with the doll's head position changed between scans. One nurse operated the scanner (Figure 4), while the other assisted in repositioning the head, mirroring routine care practices. After the three scans, the nurse that operated the scanner measured HC using measuring tape, adhering to standard protocol (Figure 5). After completing a session, nurses exchanged tasks and conducted another measurement session following the same procedure.

Merging scans to one complete head

For each scan session, a comprehensive 3D image of the head was constructed by merging the nine raw 3D images into one. The transition from raw scans to HC/CrV values is illustrated in Figure 6. The merging of all scans was undertaken collaboratively by two researchers: a KTO-student and an Vectory3 employee, who divided the nurses' scan sessions between them.

Calculation of HC and CrV

HC and CrV measurements were obtained from the 3D image using CraniumPy software (Abdel-Alim, 2023). In the calculation process, a rater manually marks three anatomical landmarks on the 3D head image: the left tragus, right tragus, and nasion (as depicted in Figures 6, 7). Subsequently, the software constructs a virtual plane passing through both tragi and nasion, computing the volume above this plane as CrV. To determine HC, the software automatically identifies the largest section above the ears, aligned parallel to the tragi-nasion plane.

To mitigate potential rater bias, HC and CrV were independently assessed from the 3D scans by three rater groups. The first rater group KTO-V3 comprised a Technical Medicine student and Vectory3 employee lacking clinical background, who evenly divided the acquired M3D scans between them. To reduce interrater variability, the two KTO-V3 raters underwent joint



FIGURE 2 A specially prepared, ventilated preterm simulation doll lies in the incubator.



FIGURE 3

Scan session setting in a NICU bed space. One nurse operated the M3D scanner using the laptop. The nurse activated the capture button on the touchscreen to start the scan. Meanwhile, the other nurse assisted in repositioning the head between scans, mimicking routine care practices.

practice sessions for marking tragi and nasion points. The second (RG) and (TS) rater 'groups' each comprised a single researcher lacking clinical background. Prior to assessment, all individual raters received informal training in landmark placement techniques.

Data analysis

Practical usability

The usability of scan sessions was assessed quantitatively by determining the percentage of successful sessions, defined as those

where HC and CrV could be accurately calculated from the comprehensive 3D head image.

Accuracy and precision

Accuracy and precision were characterized by the measurement error (ME). As depicted in Figure 8, accuracy and precision were delineated relative to the ground truth (GT), representing the true value of the measured object. The method for establishing GT is explained in the next paragraph.



FIGURE 4

The M3D Graphical User Interface is displayed on a laptop with a touchscreen. Three 'live' camera previews assist in orienting the scanner relative to the head. To position the doll's head correctly, it must be placed within the green circle by gently sliding the scanner over the incubator's cover. Before initiating a capture (using the green button), the head's orientation is selected from one of three buttons featuring graphical head contours, corresponding to the mid-camera preview. After a capture, that the selected button turns green, providing visual feedback to users regarding the completed head positions. A 3D file of the capture, labeled with a suffix indicating the head's position (front, left, right), is automatically stored on the laptop. The doll's hat, although part of the realistic setting, is removed before each scan.



The measuring tape currently utilized for measuring HC at the NICU involved in the study is the HappyMeasure tape. This tape, made of Tyvek plastic paper, was specifically developed for in-hospital use by the Erasmus MC department of Create4Care, Rotterdam, the Netherlands.

Accuracy was measured by the mean or median absolute error (MAE) if ME was normally distributed, or the median absolute deviation (MAD) if not. Precision, indicative of spread of the ME, was quantified by the 95% reference interval of the total ME range,

encompassing 95% of observed MEs. This interval was preferred over the full ME range to exclude outliers. For non-normally distributed MEs, the Q2.5 - Q97.5 95% reference interval was utilized, while normally distributed MEs were represented by the



FIGURE 6

The transition from raw scans to HC/CrV values (3 x 3 raw camera captures) into a unified head. The merging procedure uses a combination of manual and automatic algorithms, all part of the Meshlab software application software (MeshLab 2022.02, Visual Computing Lab, ISTI-CNR, Pisa, Italy). The whole procedure takes 10–20 min per head. Each individual camera capture contains all 3D surfaces (called meshes) visible for that camera. First step is 'cleaning': the head is separated from unwanted artefacts (such as the blanket, matrass, infant's shoulders) by manually selecting and deleting 3D points. This is done for all nine captures. Second step is alignment of the nine captures relative to each other, matching overlapping areas with each other. Alignment is done with the Meshlab 'align' function, using 'point based glueing'. With point based glueing, four 'matching pair of points' are manually selected on the overlapping areas of both meshes. Using the four pair of points, the algorithm roughly aligns meshes to each other in 3D space. After rough alignment, an 'iterative closest point' (ICP) algorithm automatically 'fine tunes' alignment by iteratively calculating the smallest average distances between 3D points of both meshes. After aligning two meshes, a third mesh is added and aligned with the set, etcetera, until all nine captures form one head, that is exported as one unified mesh. Finally, any minor gaps are filled to allow CrV and HC calculation by the calculation script.



and nasion.



mean \pm 1.96*SD distribution reference interval to cover 95% of MEs. Accuracy and precision were expressed both in measurement units (cm for HC, mL for CrV) and a percentage of the GT value.

The GT values were derived from accurate 3D scans from the doll's head, obtained using a medical scanner with known high accuracy (3dMD head system, purchased in 2020, 3dMD, Atlanta, USA, with a manufacturer-claimed accuracy of 0.3 mm). These head scans are 360-degree scans. HC and CrV were measured from these scans by the same measuring method as described above, utilizing the CraniumPy calculation script. Four different 3dMD-scans were measured by five raters, resulting in a total of 34 GT measurements. As there were no extreme outliers, we used the mean of these measurements as the GT values for the doll used in this study.

The influence of raters on M3D measurements was assessed by comparing the MEs among the three independent M3D rater groups. To assess the overall accuracy and precision of the M3D HC and CrV measurements, the measurements from all three rater groups were combined and analysed as a single rater group.

Accuracy and precision for clinical use

For medical measuring instruments, the required accuracy and precision depend on their intended use, with the permissible ME being contingent on the smallest clinical change the instrument should detect (ISO IOfS, 2023). For the M3D scanner, this clinical change is the weekly growth increase in HC and CrV, as this scanner is intended for weekly growth measurements (Villar et al., 2016; Health UDo, 2009; Greer and Olsen, 2013). Figure 9 demonstrates the relationship between weekly measurements and the true weekly growth using a fictive data plot.

To ensure sufficient accuracy for the total population, the clinically allowed ME should be based on the smallest weekly increases. The weekly increases can be expressed in either absolute values (in cm for HC, and mL for CrV) or proportional increases (in %). Table 1 summarises the weekly increase in HC and CrV for preterm infants. We aggregated weekly HC and CrV growth data from multiple sources, including growth reference charts and published studies, to compile this information (Vermeulen et al., 2021; Young et al., 2020; RCPCH, 2019;



Dieks et al., 2022; Cole et al., 1998). Detailed tables of the weekly increases are available in the supplementary materials.

Statistical analysis

We employed descriptive statistics to quantify accuracy and precision, along with MEs, as described above. Normality of measurement distributions and MEs was assessed using Shapiro-Wilk normality tests, supplemented by violin plots for visual representation. Rater group differences were evaluated via one sample T-tests for assumed normal distributions and Wilcoxon rank sum tests for non-normal distributions. Rater effects were examined using either one-way ANOVA, if normality was assumed, or Kruskal–Wallis tests, if assumed non-normal. To visually depict accuracy and precision, we generated point plots illustrating MEs for each rater group and across all raters.

Statistical analyses were performed using R Statistical Software Version 4.2.1 (R Foundation of Statistical Computing, Vienna, Austria), with significance set at p < 0.05 (two-tailed).

Results

Practical usability

NICU nurses conducted a total of fifty-six scan sessions, resulting in successful acquisition of complete head scans in 25 instances (success rate of 44.6%), from which scans HC and CrV could be derived.

Weekly increase HC	Absolute [cm]	Proportional	Weekly increase CrV	Absolute [mL]	Proportional
Minimum value	0.4	1.0%	Minimum value	14.3	3.2%
Maximum value	1.3	5.4%	Maximum value	44.0	14.7%
Mean	0.9	3.2%	Mean	30.5	7.6%
Median	0.9	3.3%	Median	34.0	7.1%

TABLE 1 Weekly increase of head circumference and cranial volume for preterm infants.

HC, head circumference; CrV, cranial volume.



FIGURE 10

The image on the left displays a complete M3D head scan alongside its corresponding measurement outcomes. Two small green dots are discernible, denoting the manually marked landmarks for the tragus and nasion, identified by one of the raters. On the right image, the computational software utilizes the tragi-nasion plane to trim the head with, revealing the resulting volume. A prominent red line indicated the maximum circumference above this plane.

Figure 10 illustrate a successful head scan along with its corresponding measurement outcomes. In thirty-one scan sessions, HC and CrV measurements were unsuccessful due to various reasons: a) model failure, where the back of the head was incomplete (n = 16); b) capture failure, where not all camera views per capture were saved (n = 13); c) capture failure, where the capture was accidentally overwritten by other participant (n = 1); and d) occlusion of the head, making the selection of reference marks impossible (n = 1).

Accuracy and precision of HC measurements

Statistics for HC measured values and corresponding measurement errors are presented in Table 2. Figure 11 provides a visual interpretation of the distribution of HC measurement errors for each rater. Detailed statistic tables are available in the supplementary materials. Shapiro-Wilk normality tests and violin plots indicate that the measurements are not normally distributed for both individual rater groups and the combined raters. Most M3D measured values are below GT. For all individual rater groups and the combined raters, the median HC measurements deviate significantly from the GT. The Kruskal–Wallis test further confirms that the raters' influence on median HC across all groups is significant (p= 0.003). Accuracy and precision metrics are presented in Table 2. While the measuring tape has a higher accuracy with a median absolute deviation of 0.2 cm (proportional 0.9% of GT value) compared to M3D's accuracy (combined raters) of 0.4 cm (1.5% of GT), its precision of 1.6 cm (6.3% of GT) is more than twice as low as M3D's precision of 0.7 cm (2.9% of GT).

Accuracy and precision of CrV measurements

Table 3 presents statistics for CrV measured values and CrV MEs. Figure 12 visualizes the distribution of CrV measurement errors for TABLE 2 Head circumference measurements conducted using M3D and measuring tape across different rater groups.

HC	GT = 24.8 cm	T = 24.8 cm						Accuracy ^d		Precision ^e	
Device	Rater (Group name)ª	n	Median HC [cm]	Q25 – Q75	Median ME [cm]	Wilcoxon p-value ^b	Kruskal–Wallis p-value ^c	MAD (cm)	MAD as % of GT	95% interval of ME (cm)	95% interval in % of GT
Таре	Nurses	56	25.0	24.6-25.0	0.2	0.464	NA	0.2	0.9	1.6	6.3
M3D	V3-KTO	25	24.5	24.4-24.6	-0.3	<0.001	0.003	0.3	1.2	0.6	2.6
M3D	RG	25	24.3	24.3-24.5	-0.5	<0.001	0.003	0.5	1.9	0.6	2.5
M3D	TS	25	24.3	24.3-24.4	-0.5	<0.001	0.003	0.5	1.9	0.7	2.8
M3D	Raters combined	75	24.4	24.3-24.6	-0.4	<0.001	NA	0.4	1.5	0.7	2.9

Tape, Measuring Tape; M3D, MONITOR3D; GT, ground truth; HC, head circumference; ME, Measurement Error= measured value–GT, negative value = below GT; 95% interval = Q97.5 - Q2.5; MAD, Medium Absolute Deviation from GT, Medium Absolute ME.

^aRater groups. Tape = Measuring tape used by nurses (n=25 raters).

V3-KTO, M3D scan + CraniumPy, by researchers (divided over 2 raters).

RG, M3D scan + CraniumPy, by researcher (1 rater).

TS, M3D scan + CraniumPy, by researcher (1 rater).

Combined raters =V3-KTO & RG & TS.

^bWilcoxon-signed rank test with continuity correction, difference HC-GT.

^cKruskal Wallis rank-based test applied to assess differences in medians among the three M3D rater groups.

^dSince head circumference measurement errors are not normally distributed, accuracy is expressed as the median absolute deviation from the ground truth.

^ePrecision is represented by the 95% interval of observed errors.



Plot of head circumference measurement errors for each rater group, with bee-swarmed points and violin plot combined. The Error Interval, as shown in the plot, acts as measure of precision. This error interval is the range between the Q2.5 and Q97.5 limits, encompassing 95% of observed errors. 'M3D, combined' represents the overall performance of the M3D scanner.

each rater. Shapiro-Wilk normality tests and violin plots indicate that the measurements are normally distributed for all individual rater groups and the combined raters. Most M3D measured values are below GT. For all individual rater groups and the combined raters, the mean CrV measurements deviate significantly from the GT. The one-way ANOVA further confirms that the raters' influence on mean CrV across all groups is significant (p= 0.003).

Discussion

The current study evaluated the usability, accuracy and precision of head circumference (HC) and cranial volume (CrV) measurements using the novel M3D 3D-scanner in a simulated NICU setting. Nurses performed scan sessions with the M3D scanner by scanning the head of a preterm doll. Additionally, they measured HC using the existing measuring tape. Researchers derived HC and CrV from the M3D 3D scans. Regarding usability, integrating M3D scan sessions with routine care handling is feasible, although the scan success rate of 44.6% is considered too low. The accuracy and precision of the M3D scanner are clinically sufficient for weekly measurements in 80% of the population for HC and 70% for CrV among preterm infants, based on the growth data sources used in this study. Notably, the currently used measuring tape does not meet the clinical standards set in this study. Detailed findings are discussed below.

Practical usability

Integrating scan sessions into routine care

It proved feasible to integrate M3D scan sessions into routine care moments of ELBW infants in incubators. Utilizing the

TABLE 3 Cranial volume measurements using M3D by different rater groups.

CrV	GT = 209.9 mL								Accuracy ^d		Precision ^e	
Device	Rater (group name)ª	n	Mean CrV [mL]	SD	Mean ME [mL]	SD ME	t-test p-value⁵	ANOVA p-value ^c	MAE (mL)	MAE in % of GT	±1.96*SD 95% interval ME (mL)	95% interval in % of GT
M3D	V3-KTO	25	204.3	5.2	-5.6	5.2	<0.001	0.0032	6.5	3.1	20.5	9.8
M3D	RG	25	204.0	5.3	-5.9	5.3	<0.001	0.0032	6.7	3.2	20.7	9.9
M3D	TS	25	199.5	5.7	-10.4	5.7	< 0.001	0.0032	10.7	5.1	22.2	10.6
M3D	Raters combined	75	202.6	5.8	-7.3	5.8	<0.001	NA	8.0	3.8	22.6	10.8

M3D,MONITOR3D; GT, ground truth; CrV, cranium volume; SD, standard deviation; ME, Measurement Error= measured value–GT, negative value = below GT; MAE, Mean Absolute Error from GT, Mean Absolute ME.

^aRater groups.

V3-KTO, M3D scan + CraniumPy, by researchers (divided over n=2 raters).

RG, M3D scan + CraniumPy, by researcher (n=1 rater).

TS, M3D scan + CraniumPy, by researcher (n=1 rater).

Raters combined =V3-KTO and RG and TS.

^bOne-sample *t*-test: mean CrV-GT.

^cAnalysis of Variance test for difference in means between the three M3D rater groups.

^dAs the distribution is normal, accuracy is expressed as the mean absolute error from GT.

^ePrecision is represented by the 95% interval, calculated as ±1.96*SD.

established protocol for routine care head position changes, the scan sessions do not introduce additional disturbances to those already inherent in routine care. The short 30 msec capture time only briefly interrupts routine care procedures. Needed three captures can be obtained within two of maximum three successive care moments, during the same day. This integration stands out as a significant advantage of the M3D scanner. Moreover, feedback from the majority of nurses indicated that the M3D scanner is less intrusive for infants compared to the currently used measuring tape. However, assessing the feasibility was somewhat limited by the use of a doll model instead of real patients. The material and stiffness of the doll made the simulation of head repositioning somewhat unnatural.

CrV as longitudinal preterm growth parameter

Potentially the most significant clinical advantage of the M3D's measuring technique, if successfully integrated in NICU care, is adding CrV as longitudinal growth parameter for ELBW infants. Several studies address the clinical added value of 3D volumetric growth data relative to traditional 2D data such as HC (Vermeulen et al., 2021; Burkhardt et al., 2019; Santander et al., 2020; Martini et al., 2018; Meyer-Marcotty et al., 2018; Ifflaender et al., 2013). In relation to this, skull deformations, a relatively common issue in preterm infants, can distort HC measurements, leading to inaccurate assessments. As deformations influence HC, but not skull volume, this also speaks for CrV as more reliable longitudinal growth parameter than HC. Skull deformities such as plagiocephaly can be caused by a prolonged head position, despite protocolized head position changes every three to 4 hours (Kubik et al., 2024; De Koning, 2020). In addition, deformations can also result from

continuous positive ear pressure (CPAP) breathing systems, attached with tight straps or hats.

Few other studies have explored feasible techniques for measuring preterm CrV. Dieks et al. (Dieks et al., 2022) conducted longitudinal CrV measurements on 26 preterm infants from 28 up to 40 weeks GA using a handheld 3D camera, suggesting its feasibility for routine neonatal care (Dieks et al., 2022). Similarly, a literature review identified the same handheld 3D camera as suitable for stress-free body size measurements of extremely preterm infants in incubators (Van Gils et al., 2022). However, this method requires eye protectors and a nylon cap on the infant's head, along with captures from ten different viewing angles, necessitating removal of the incubator's cover and repositioning the infant, or holding the infant in a parent's arms (Santander et al., 2021). Vermeulen et al. measured CrV growth using a tabletop 3D scanner, but this approach necessitated placing infants completely outside the incubator (Vermeulen et al., 2021). This would be extremely disruptive and potentially hazardous for preterm ELBW infants due to risks of hypothermia or accidental extubation.

Success rate of scan sessions

The success rate of scan sessions in our study was 44.6% (25 out of 56 sessions), which we consider too low for clinical use. In comparison, Andrews et al., 2019, in a similar study measuring HC and length of preterm infants with a 3D camera, reported a success rate of 76% (Andrews et al., 2018). One suspected reason for the high failure rate in our study is the lack of immediate feedback (the resulting 3D image) following a scan. Such a 'preview' could have aided in assessing scan success, for example, in cases where the view may have been obstructed by artefacts such as endotracheal tubes. Additionally, the cameras' limited field of view created 'blind



Plot of cranial volume measurement errors for each rater group, with bee-swarmed points and violin plots combined. The Error Interval, as shown in the plot, acts as measure of precision. This error interval is the range between the \pm 1.96*SD limits, encompassing 95% of the normal distribution. 'M3D, combined' represents the overall performance of the M3D scanner.

spots', leading to missing 3D data and hindering the construction of a comprehensive head image.

Accuracy and precision

Accuracy and precision for clinical use

In assessing clinical suitability, precision holds greater importance than accuracy. The 95% ME reference interval, indicative of precision, should be notably narrower than the anticipated weekly growth rate. The impact of ME interval size relative to growth is elucidated through four plots with fictive data in the supplementary materials. For this study, we propose that to ensure adequate clinical accuracy and precision across the entire preterm population, the measuring device's ME interval must be equal to or smaller than the smallest anticipated weekly growth. This will prevent the event that negative growth (shrinkage) is measured by mistake, when in truth, the infant is growing according to its reference curve. This smallest weekly growth was determined to be 0.4 cm (1.0% proportional) for HC and 14.3 mL (3.2% proportional) for CrV.

Determining the optimal size of the error interval relative to weekly remains somewhat arbitrary. Currently, there are no comparable studies addressing this issue. However, one study that measured preterm length and HC using 3D technology interpreted the error interval as indicative of clinical usability. The authors determine the 95% limits of agreement to be +2.1 to -1.8 cm, stating that "...such a difference is likely to be clinically important, leading to mis-allocation of the correct growth percentile, especially in smaller infants" (Andrews et al., 2019).

It could be argued that the error interval should not exceed the minimum expected growth, preventing the possibility of measuring

'negative' growth when normal growth is occurring. Following this rationale, the M3D scanner could measure weekly HC growth with sufficient accuracy for approximately 80% of preterm infants, while weekly CrV could be assessed for 70% of preterm infants: the M3D's HC error interval of 0.7 cm approximately matches the 80s percentile of HC weekly growth distribution, and CrV precision of 22.6 mL approximately the 70s percentile of CrV weekly growth. It should be noted that these HC/CrV weekly growth distributions were limited to the aggregated data sources identified for this study. The HC growth data, with sufficient data sources identified, can be

considered representative for a large part of total preterm population. However, for preterm CrV data, only two data sources could be identified, with relative small sample sizes.

Head circumference with existing measuring tape *versus* M3D

The precision of the existing measuring tape appears insufficient for weekly HC measurements, as indicated by the data in this study. This conclusion stems from the measuring tape's ME interval of 1.6 cm (equivalent to 6.3%) significantly exceeds the anticipated weekly HC growth range of 0.4–1.3 cm. Despite the measuring tape demonstrating better accuracy, indicated by its lower median absolute error compared to the M3D scanner, the scanner's superior precision makes it the preferable option.

However, it is worth noting a potential bias in the M3D scanner's accuracy due to a systematic error, discussed in detail below. As a limitation in HC measurements, we note that the HC measurements with 3D scans, including the HC GT measurement, measured (the largest section) parallel to the tragi-nasion plane, where the manual measuring tape, as stipulated in the NICU's protocol for HC measurements, measured at the largest circumference 'above eyebrows and ears'. This might have caused small bias when comparing measuring tape measurements with GT.

Raters' influence on measurements

The raters in this study had a significant influence on the accuracy for both HC and CrV measured using the M3D scanner. Because landmarks were manually placed, significant interrater variability was anticipated. However, the plots show that interrater precision for M3D measurements, represented by the spread of M3D ME, hardly differed between raters. In contrast, the intrarater variability was larger for CrV than for HC, indicating that CrV measurements are less precise than HC measurements. This is likely because the calculation script is more sensitive to user input for CrV than for HC. For HC, the exact position of the traginasion plane has a minor influence, because the script 'searches' for the largest section above the ears parallel to this plane. In contrast, for CrV the location of landmarks directly affects the volume calculation above the plane. This effect is further enhanced by the small size of the doll's head and the cubic nature of volume measurements, where small differences in landmark positions result in relatively large CrV differences between raters. Therefore, this technique makes it is harder to accurately measure CrV than HC. Several studies have used marking of the nasion and tragi for deriving HC and CrV from 3D scans and noted user influence on the measurements (Burkhardt et al., 2019; Ifflaender et al., 2013; Barbero-García et al., 2017; Geil and Smith, 2008; Weinberg et al., 2006). Tareq Abdel et al. developed CraniumPy software (Abdel-Alim et al., 2023) for assessing head morphology and size, validating its intra- and interrater reliability of the method with full 360-degree, highly detailed scans of infants, though not preterm infants. Their study showed that manual marking of landmarks introduces an intra- and interobserver variability of 3%. Ifflaender et al. measured HC and CrV of preterm infants and found high agreement between raters, with mean difference between raters of -0.1 ± 0.74 mm ($-0.005\% \pm 0.2\%$) for HC and 8.1 \pm 6.5 mL (1.5% \pm 1.1%) for CrV (Ifflaender et al., 2013).

Our primary limitation for accuracy and precision was the lack of morphological details in the 3D images from the M3D scanner, making it difficult to distinguish anatomical landmarks such as the tragus and nasion. The low resolution of the depth cameras was especially problematic around the ears, complicating the exact location of the tragus. Additionally, fixation tape used to secure the feeding and breathing tubes in the nose, covered the nasion, making it more difficult to locate. The absence of colour in the 3D images further hindered the identification of the anatomical points, as CraniumPy only provided images in grey. The lack of detail in the raw 3D images also made it harder to merge the three captures into an accurate and comprehensive head image.

Suspected systematic error

A suspected systematic technical error may have further influenced inaccuracy. This suspicion arises from the observation that the mean measurements from all individual rater groups were consistently below GT. We believe the cause lies in the complex chain of actions from 3D scans to the final measurements.

Comparing measurements against ground truth values

Using a doll instead of real patients allowed for the comparison of measured values against accurately defined GT values, which we consider a strength of this study. Most studies assessing 3D body size measuring techniques compare the new method's performance relative to an existing method (Van Gils et al., 2022). However, obtaining accurate GT values for 3D body size measures such as CrV in real patients is often too complex, with CT or MRI too invasive or expensive. Comparing a novel device only against the traditional 'gold standard' technique is debatable: a perfect correlation does not necessarily indicate how close the measured values are to the true GT values. In fact, the new technique's measurements may be closer to the true values than those obtained with the traditional 'gold standard' instrument due to unknown limitations of the latter. A limitation of our study is that it is impossible to define GT with 100% accuracy. The clinical 3dMD head system scanner we used had an accuracy of 0.3 mm, as stated by the manufacturer. Additionally, there is interrater variability in placing landmarks in CraniumPy (Abdel-Alim et al., 2023).

Future studies

Future studies should focus on improving M3D's success rate of scans and accuracy and precision. Such a technically improved M3D scanner should first be validated in a simulated setting to confirm accuracy and precision before being validated with real patients. Such study with real patients could collect the first-in-the-world longitudinal preterm CrV data of infants <28 weeks GA. The most effective step to further improve the accuracy and precision of the M3D scanner would be more accurate marking of anatomical points, requiring 3D images with better morphological details and added colour information. Potential systematic errors should be analyzed and addressed under controlled lab conditions. To enhance the success rate of scan sessions, the 3D image should be displayed directly on screen after a scan, allowing users to verify if a capture was successful or needs to be redone. Ideally, the M3D measuring technique should be expanded to include body length in addition to HC and CrV, creating an all-in-one device for stress-free growth monitoring of ELBW infants. In addition, with improved scan quality, the scope of the M3D could be further extended to morphological head analysis, such as monitoring of preterm skull deformations, for preterm infants that cannot be scanned outside their incubator.

Conclusion

The M3D measuring technique is suitable for longitudinal HC and CrV growth monitoring of ELBW infants in incubators. Scan sessions can be integrated into routine care, without introducing additional disturbances to the infant. However, its current scan success rate is too low for practical usability. M3D's current accuracy and precision are clinically sufficient for weekly measurements, for approximately 70% of preterm infants for CrV, and 80% for HC. The precision of the existing measuring tape is inadequate for any preterm infant. Consequently, M3D already presents a better alternative for measuring head circumference among ELBW infants in incubators, with no disturbance to the infant.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Clinical Technological Research – Human Research Ethics Committee (KTO-HREC) of Delft University of Technology, Application number KTO-HREC21005. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

RG: Writing-original draft, Writing-review and editing. OH: Writing-review and editing. RK: Writing-review and editing. IR: Writing-review and editing. JD: Writing-review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. RG's contribution to this study is partly financed by the Dutch Research Council (NWO) with a Doctoral Grant for Teachers with project number 023.012.013.

Acknowledgments

We thank Vectory3 for their device development efforts and support during validation. We thank all involved software engineering students of Rotterdam University of Applied Sciences. We thank Thomas Feijts, Crystal Wang, Bram van Deurzen and Hil Steketee, students of the Delft University of Technology at the time of this study, for preparing and executing the validation. We thank all NICU nurses involved, with special thanks to Megan Grotenhuis and Myrthe Canter-Visscher. We thank Tareq Abdel-Alim for supporting us with CraniumPy, Timothy Singowikromo for contributing to data analysis and general support during this research, and Ko Hagoort for text editing.

Conflict of interest

At the time of the study, Erasmus MC-Create4Care had a partnership agreement with Vectory3 for co-developing a 3D scanner for measuring HC and CrV. This agreement covered intellectual property claims, input of front and background knowledge, and development efforts.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abdel-Alim, T. (2023). T-AbdelAlim/CraniumPy: CraniumPy v0.4.2 (v0.4.2). Computer software. Zenodo. Available at: https://doi.org/10.5281/zenodo.8310267

Abdel-Alim, T., Tio, P., Kurniawan, M., Mathijssen, I., Dirven, C., Niessen, W., et al. (2023). Reliability and agreement of automated head measurements from 3-dimensional photogrammetry in Young children. *J. craniofacial Surg.* 34, 1629–1634. doi:10.1097/scs.00000000009448

Als, H., Duffy, F. H., McAnulty, G., Butler, S. C., Lightbody, L., Kosta, S., et al. (2012). NIDCAP improves brain function and structure in preterm infants with severe intrauterine growth restriction. *J. Perinatol.* 32 (10), 797–803. doi:10.1038/jp.2011.201

Andrews, E., Ashton, J. J., Pearson, F., Beattie, R. M., and Johnson, M. J. (2018). "Using handheld 3D scanning as a technique for measuring neonatal head circumference and length," in *PIER Conference 2018 - Paediatric innovation*, *education & research network*. Wessex, England.

Andrews, E. T., Ashton, J. J., Pearson, F., Beattie, R. M., and Johnson, M. J. (2019). Handheld 3D scanning as a minimally invasive measuring technique for neonatal anthropometry Handheld 3D scanning as a minimally invasive measuring technique for neonatal anthropometry, *Clin. Nutr. ESPEN*, 33:279–282. doi:10.1016/j.clnesp.2019.06.012

Barbero-García, I., Lerma, J. L., Marqués-Mateu, Á., and Miranda, P. (2017). Low-cost smartphone-based photogrammetry for the analysis of cranial deformation in infants. *World Neurosurg.* 102, 545–554. doi:10.1016/j.wneu.2017.03.015

Brummelte, S., Grunau, R. E., Chau, V., Poskitt, K. J., Brant, R., Vinall, J., et al. (2012). Procedural pain and brain development in premature newborns. *Ann. Neurology* 71 (3), 385–396. doi:10.1002/ana.22267

Burkhardt, W., Schneider, D., Hahn, G., Konstantelos, D., Maas, H. G., and Rüdiger, M. (2019). Non-invasive estimation of brain-volume in infants. *Early Hum. Dev.* 132, 52–57. doi:10.1016/j.earlhumdev.2019.03.020

Cole, T. J., Freeman, J. V., and Preece, M. A. (1998). British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. *Statistics Med.* 17 (4), 407–429. doi:10.1002/(sici)1097-0258(19980228)17:4<407::aid-sim742>3.3.co;2-c

De Koning, M. (2020). Protocol: wisselligging bij neonaat met bewegingsarmoede (Body and head positioning management in neonates with reduced mobility, nursing protocol). Quality Management System, Erasmus MC University Medical Centre Rotterdam 2020.

Dieks, J. K., Jünemann, L., Hensel, K. O., Bergmann, C., Schmidt, S., Quast, A., et al. (2022). Stereophotogrammetry can feasibly assess 'physiological' longitudinal threedimensional head development of very preterm infants from birth to term. *Sci. Rep.* 12 (1), 8940. doi:10.1038/s41598-022-12887-x

Doesburg, S. M., Chau, C. M., Cheung, T. P. L., Moiseev, A., Ribary, U., Herdman, A. T., et al. (2013). Neonatal pain-related stress, functional cortical activity and visual-perceptual abilities in school-age children born at extremely low gestational age. *Pain* 154 (10), 1946–1952. doi:10.1016/j.pain.2013.04.009

Geil, M. D., and Smith, A. (2008). Accuracy and reliability of a system for the digital capture of infant head shapes in the treatment of cranial deformities. *JPO J. Prosthetics Orthot.* 20 (2), 35–38. doi:10.1097/jpo.0b013e318169c439

Greer, F. R., and Olsen, I. E. (2013). How fast should the preterm infant grow? *Curr. Pediatr. Rep.* 1 (4), 240–246. doi:10.1007/s40124-013-0029-1

Grunau, R. E. (2013). Neonatal pain in very preterm infants: long-term effects on brain, neurodevelopment and pain reactivity. *Rambam Maimonides Med. J.* 4 (4), e0025. doi:10.5041/rmmj.10132

Grunau, R. E., Whitfield, M. F., Petrie-Thomas, J., Synnes, A. R., Cepeda, I. L., Keidar, A., et al. (2009). Neonatal pain, parenting stress and interaction, in relation to cognitive and motor development at 8 and 18 months in preterm infants. *Pain* 143 (1-2), 138–146. doi:10.1016/j.pain.2009.02.014

Health UDo (2009). UK-WHO growth charts - neonatal and infant close monitoring. Ifflaender, S., Rüdiger, M., Koch, A., and Burkhardt, W. (2013). Three-Dimensional digital capture of head size in neonates - a method evaluation. *PLoS ONE* 8 (4), e61274. doi:10.1371/journal.pone.0061274

ISO IOfS (2023). Accuracy of measurement methods and results—part 1: general principles and definitions. 2 ed. Geneva, Switzerland: International Organization for Standardization.

Kubik, A., Welton, J., Holmes, L., Struwe, L., and Gonzales, K. (2024). The PLAY bundle for head deformities: a NICU quality improvement project. *J. Neonatal Nurs.* 30, 263–270. doi:10.1016/j.jnn.2023.10.006

Martini, M., Klausing, A., Lüchters, G., Heim, N., and Messing-Jünger, M. (2018). Head circumference - a useful single parameter for skull volume development in cranial growth analysis? *Head. Face Med.* 14 (1), 3. doi:10.1186/s13005-017-0159-8

Meyer-Marcotty, P., Kunz, F., Schweitzer, T., Wachter, B., Bohm, H., Waßmuth, N., et al. (2018). Cranial growth in infants-A longitudinal three-dimensional analysis of the first months of life. *J. Cranio-Maxillofacial Surg.* 46 (6), 987–993. doi:10.1016/j.jcms. 2018.04.009

Moody, C., Callahan, T. J., Aldrich, H., Gance-Cleveland, B., and Sables-Baus, S. (2017). Early initiation of newborn individualized developmental care and assessment program (NIDCAP) reduces length of stay: a quality improvement project. *J. Pediatr. Nurs.* 32, 59–63. doi:10.1016/j.pedn.2016.11.001

Ranger, M., Chau, C. M., Garg, A., Woodward, T. S., Beg, M. F., Bjornson, B., et al. (2013). Neonatal pain-related stress predicts cortical thickness at age 7 years in children born very preterm. *PLoS One* 8 (10), e76702. doi:10.1371/journal.pone. 0076702

Ranger, M., and Grunau, R. E. (2014). Early repetitive pain in preterm infants in relation to the developing brain. *Pain Manag.* 4 (1), 57–67. doi:10.2217/pmt.13.61

Ranger, M., Zwicker, J. G., Chau, C. M., Park, M. T., Chakravarthy, M. M., Poskitt, K., et al. (2015). Neonatal pain and infection relate to smaller cerebellum in very preterm children at school age. *J. Pediatr.* 167 (2), 292–298.e1. doi:10.1016/j.jpeds.2015.04.055

RCPCH (2019). UK-who growth charts-neonatal and infant close monitoring (nicm) london. UK: Royal College of Paediatrics and Child Health.

Santander, P., Quast, A., Hubbert, J., Horn, S., Meyer-Marcotty, P., Kuster, H., et al. (2020). Three-dimensional head shape acquisition in preterm infants - translating an orthodontic imaging procedure into neonatal care. *Early Hum. Dev.* 140, 104908. doi:10. 1016/j.earlhumdev.2019.104908

Santander, P., Quast, A., Hubbert, J., Juenemann, L., Horn, S., Hensel, K. O., et al. (2021). Stereophotogrammetric head shape assessment in neonates is feasible and can identify distinct differences between term-born and very preterm infants at term equivalent age. *Sci. Rep.* 11 (1), 21155. doi:10.1038/s41598-021-00680-1

VandenBerg, K. A. (2007). Individualized developmental care for high risk newborns in the NICU: a practice guideline. *Early Hum. Dev.* 83 (7), 433–442. doi:10.1016/j. earlhumdev.2007.03.008

Van Gils, R. H. J., Wauben, LSGL, and Helder, O. K. (2022). Body size measuring techniques enabling stress-free growth monitoring of extreme preterm infants inside incubators: a systematic review. *PLOS ONE* 17 (4), e0267285. doi:10.1371/journal.pone. 0267285

Vermeulen, M. J., Burkhardt, W., Fritze, A., Roelants, J., Mense, L., Willemsen, S., et al. (2021). Reference charts for neonatal cranial volume based on 3D laser scanning to monitor head growth. *Front. Pediatr.* 9, 654112. doi:10.3389/fped.2021.654112

Villar, J., Altman, D. G., Purwar, M., Noble, J. A., Knight, H. E., Ruyan, P., et al. (2013). The objectives, design and implementation of the INTERGROWTH-21st Project. *Bjog* 120 (Suppl. 2), 9–26. doi:10.1111/1471-0528.12047

Villar, J., Giuliani, F., Bhutta, Z. A., Bertino, E., Ohuma, E. O., Ismail, L. C., et al. (2015). Postnatal growth standards for preterm infants: the preterm postnatal follow-up study of the INTERGROWTH-21(st) project. *Lancet Glob. Health* 3 (11), e681–e691. doi:10.1016/s2214-109x(15)00163-1

Villar, J., Giuliani, F., Fenton, T. R., Ohuma, E. O., Ismail, L. C., and Kennedy, S. H. (2016). INTERGROWTH-21st very preterm size at birth reference charts. *Lancet* 387 (10021), 844–845. doi:10.1016/s0140-6736(16)00384-6

Weinberg, S. M., Naidoo, S., Govier, D. P., Martin, R. A., Kane, A. A., and Marazita, M. L. (2006). Anthropometric precision and accuracy of digital three-dimensional photogrammetry: comparing the Genex and 3dMD imaging systems with one another and with direct anthropometry. *J. Craniofacial Surg.* 17 (3), 477–483. doi:10.1097/00001665-200605000-00015

Young, A., Andrews, E. T., Ashton, J. J., Pearson, F., Beattie, R. M., and Johnson, M. J. (2020). Generating longitudinal growth charts from preterm infants fed to current recommendations. *Archives Dis. Child.* - *Fetal Neonatal Ed.* 105 (6), 646–651. doi:10. 1136/archdischild-2019-318404

Zevenbergen, A. (2020). MONITOR3D: design of a 3D scanner for measuring preterm infant head circumference. Master thesis integrated product design. Delft, Netherlands: Faculty of Industrial Design Engineering, Delft University of Technology. available at: https://repository.tudelft.nl/record/uuid:a59c476e-ca35-438d-9c8b-485eacdb9d72.