American Cleft Palate–Craniofacial Association

Original Article

The Cleft Palate-Craniofacial Journal
1-6

DOI: 10.1177/10556656211061021
journals.sagepub.com/home/cpc
SAGE

"Validation of Artificial Intelligence Severity Assessment in Metopic Craniosynostosis"

Alexandra Junn, AB[1], Jacob Dinis, BS[1], Sacha C. Hauc, BS, MPH[1], Madeleine K. Bruce, BA[2], Kitae E. Park, MD[3], Wenzheng Tao[4], Cameron Christensen, MS[4], Ross Whitaker, PhD[4], Jesse A. Goldstein, MD[2], and Michael Alperovich, MD, MSc[1]

Abstract

Objective: Several severity metrics have been developed for metopic craniosynostosis, including a recent machine learning-derived algorithm. This study assessed the diagnostic concordance between machine learning and previously published severity indices.

Design: Preoperative computed tomography (CT) scans of patients who underwent surgical correction of metopic craniosynostosis were quantitatively analyzed for severity. Each scan was manually measured to derive manual severity scores and also received a scaled metopic severity score (MSS) assigned by the machine learning algorithm. Regression analysis was used to correlate manually captured measurements to MSS. ROC analysis was performed for each severity metric and were compared to how accurately they distinguished cases of metopic synostosis from controls.

Results: In total, 194 CT scans were analyzed, 167 with metopic synostosis and 27 controls. The mean scaled MSS for the patients with metopic was 6.18 ± 2.53 compared to 0.60 ± 1.25 for controls. Multivariable regression analyses yielded an R-square of 0.66, with significant manual measurements of endocranial bifrontal angle (EBA) ($P = 0.023$), posterior angle of the anterior cranial fossa ($p < 0.001$), temporal depression angle ($P = 0.042$), age ($P < 0.001$), biparietal distance ($P < 0.001$), interdacryon distance ($P = 0.033$), and orbital width ($P < 0.001$). ROC analysis demonstrated a high diagnostic value of the MSS (AUC = 0.96, $P < 0.001$), which was comparable to other validated indices including the adjusted EBA (AUC = 0.98), EBA (AUC = 0.97), and biparietal/bitemporal ratio (AUC = 0.95).

Conclusions: The machine learning algorithm offers an objective assessment of morphologic severity that provides a reliable composite impression of severity. The generated score is comparable to other severity indices in ability to distinguish cases of metopic synostosis from controls.

Keywords

craniosynostoses, machine learning, algorithms, cephalometry

Introduction

Comprising up to 25% of non-syndromic craniosynostosis cases, metopic suture craniosynostosis can result in trigonocephaly, orbital hypotelorism, bitemporal narrowing, and deformities of the orbital rims.(Posnick et al., 1994; Kolar, 2011; Birgfeld et al., 2013) Fusion of the suture can range from mild ridging to a pronounced "keel-shaped" deformity. As physiologic closure of the metopic suture occurs between 3 and 10 months of age, however, the indication for surgical intervention for metopic synostosis remains controversial (Vu et al., 2001).

While immediate surgical correction is recommended in cases of severe deformity and observation alone is usually

[1] Department of Surgery, Division of Plastic Surgery, Yale School of Medicine, New Haven, CT, USA
[2] Department of Plastic Surgery, University of Pittsburgh Medical Center, Pittsburgh, PA, USA
[3] Department of Plastic and Reconstructive Surgery, Johns Hopkins Hospital; Baltimore, MD, USA
[4] School of Computing, University of Utah, Salt Lake City, UT, USA
*Alexandra Junn, Jacob Dinis, Co-irst Author

Corresponding Author:
Michael Alperovich, Yale School of Medicine, Department of Surgery, Division of Plastic Surgery, 330 Cedar Street, Boardman Building, 3rd Floor, New Haven, CT 06510 USA.
Email: michael.alperovich@yale.edu

appropriate for most mild cases, the decision of whether or not to intervene in moderate cases is often equivocal. Previous studies have sought to quantify the severity of deformity in metopic synostosis to provide insight into defining a threshold for intervention. Various indices composed of linear distances, angles, and ratios based on imaging have been suggested.(Posnick et al., 1994; Bottero et al., 1998; Beckett et al., 2012; Kellogg et al., 2012; Ezaldein et al., 2014; Wang et al., 2016; Naran et al., 2017; Cronin et al., 2020; Chandler et al., 2021) More recently, a machine learning-derived analysis of skull shape has been used to generate an objective score measuring head shape deformity.(Bhalodia et al., 2020) Initial validation of this algorithm has demonstrated high degree of correlation between the algorithm scores and expert ratings of severity on aggregate.

This study sought to investigate the relationship between the manual severity measures and the machine learning algorithm score to assess the validity of the algorithm against the previously published metrics. In addition, the manual measurements reproduced from the literature were compared with the machine learning score on their respective abilities to accurately distinguish cases of metopic synostosis from unaffected controls.

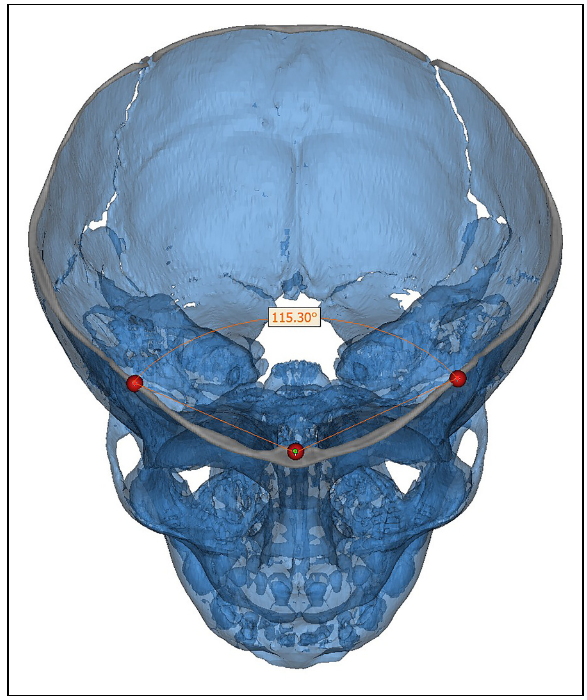

Figure 1. Endocranial bifrontal angle measurement on superior view of the skull.

Methods

This study was conducted according to the World Medical Association Declaration of Helsinki. Approval for this study was obtained by the Yale University Institutional Review Board (#2000026415), and patient consent was not obtained as only de-identified patient data was used.

Morphologic Severity Quantification

Preoperative computed tomography (CT) scans of patients who underwent surgical correction of metopic craniosynostosis were obtained from multiple US craniofacial centers as part of the Wake Forest Craniofacial Imaging Database consortium. Additionally, CT scans of control subjects who presented for trauma to Yale New Haven Hospital without any confounding craniofacial conditions were obtained.

Manual measurements were performed by manipulating Digital Imaging and Communications in Medicine data using Materialise software (Mimics version 22, 3-matic version 13, Leuven, Belgium). Measurable metopic severity indices included: endocranial bifrontal angle (EBA), adjusted EBA, frontal angle, posterior angle of the anterior cranial fossa, metopic index, horizontal cone angle, temporal depression angle, orbital rim angle, foramen ovale midline distance, and the bitemporal/biparietal distance ratio and were previously published (Supplementary Table 1; Figure 1). (Posnick et al., 1994; Bottero et al., 1998; Beckett et al., 2012; Kellogg et al., 2012; Ezaldein et al., 2014; Wang et al., 2016; Naran et al., 2017; Cronin et al., 2020; Chandler et al., 2021) Each measurement was performed twice by two independent reviewers, with each set of measurements at least 2 weeks apart. Both inter- and intra-rater reliability were confirmed with a Pearson correlation coefficient >0.9.

Independently, an unsupervised machine learning algorithm also assigned each of the CT scans a quantitative shape severity score. The scaled metopic severity score (MSS) produced by the algorithm represents the severity of the patient's head shape based on characteristic features of metopic synostosis (Figure 2). To generate the severity score, machine learning was used to combine the expert ratings from 18 craniofacial surgeons with 3-D skull-shape analysis using Shapeworks software as previously described.(Cates et al., 2017; Bhalodia et al., 2020)

Statistical Analysis

The correlations between each manual measurement of traditional metopic severity index and the computer generated MSS were analyzed through univariable linear regressions. **Those variables which were found to be independently significant on univariable regression** were subsequently included in a multivariable linear regression model to identify which morphologic severity indices had the greatest impact on the MSS. To properly evaluate relative manual measurements, age (months) and sex were controlled for within the multivariate model. Additionally, a Shapiro-Wilk-W test was utilized to ensure that the variables were all normally distributed. The requirement of homoscedasticity was verified within the examined variables by running a Lagrange multiplier test.

In cases where the measurement was a ratio, the numerator and denominator of the ratio were assessed separately. The

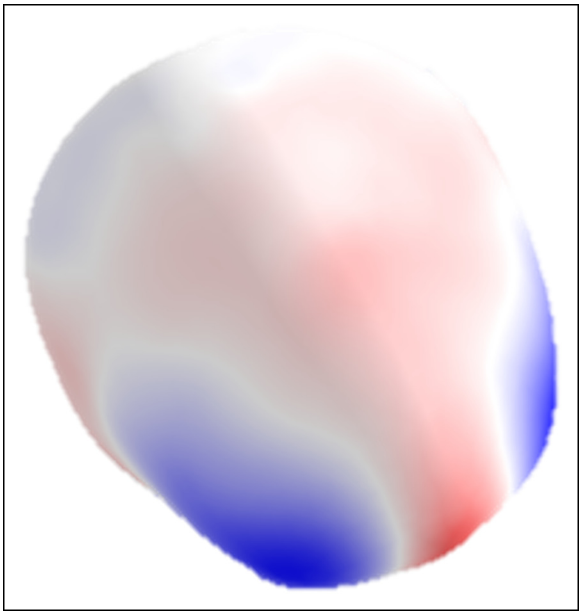

Figure 2. Heat map representing the head shape deformities expected with metopic synostosis. The red shade indicates the area where outward deformation will increase the estimated severity. The blue shade denotes the area where inward deformation will increase the estimated severity.

effects of multicollinearity within each of the models was monitored by ensuring that the variance inflation factor (VIF) was below the accepted threshold of five. In cases of collinearity, the manual measurement that generated a lower r-squared value was removed from the model. A standardized beta coefficient was obtained for each variable to evaluate the individual predictive value of each variable. To assess the fit of the model, we report adjusted R2 values.

A receiver operator characteristic curve (ROC) was also created for each severity index to assess the diagnostic value of each of the manual measurements and the MSS. Upon creation of independent receiver operating characteristic curves, the diagnostic utilities of each severity index were compared by their area under the curve (AUC). Statistical analysis was conducted using SPSS Statistics (version 25, Armonk, NY). All p-values within the model were two-tailed, and significance was set at $p < 0.05$

Results

A total of 194 CT scans were included in this study, includ-ing167 metopic craniosynostosis and 27 controls. Of the metopic synostosis cohort, 128 (76.6%) were male with a mean age of 7.18 ± 4.70 months. The mean scaled MSS for the patients with metopic was 6.18 ± 2.53. The control group of 27 CT scans was composed of 14 males (51.2%), and had a mean age of 9.21 ± 7.82 months. **There was no statistically significant difference between the ages of the metopic and control cohorts ($P = 0.136$).** The mean scaled MSS for the controls was 0.60 ± 1.25 (Table 1).

Table 1. Patient Demographics.

	Metopic (n = 167)	Control (n = 27)
Gender		
Male	128	14
Female	26	13
Age (months)	7.18 ± 4.70	9.21 ± 7.82
Scaled MSS	6.18 ± 2.53	0.60 ± 1.25

Correlation Between Machine Learning and Manual Measurements

The independent craniofacial variables selected for within the final model were: adjusted EBA, frontal coronal angles, posterior angle of the anterior fossa, orbital rim angle, temporal depression angle, biparietal distance, interdacryon distance and the upper orbital width. On multivariable regression analysis, adjusted EBA ($P = 0.023$), posterior angle of the anterior cranial fossa ($p < 0.001$), temporal depression angle ($P = 0.042$), age ($P < 0.001$), biparietal distance ($P < 0.001$), interdacryon distance ($P = 0.033$), and orbital width ($P < 0.001$) were found to be significantly related to scaled MSS. EBA was excluded due to high collinearity with adjusted EBA, with the model using adjusted EBA yielding a higher R-square of 0.66. Of the significant variables, age, biparietal distance, and orbital width had the greatest impact on the scaled MSS with standardized B coefficients of −0.37, 0.57, and −0.31 respectively (Table 2).

Comparison of Severity Indices in Distinguishing Metopic Synostosis from Controls

The diagnostic utility of the manually-determined severity indices and the algorithm-derived severity score was assessed through ROC curve analysis. Where a higher AUC indicated higher sensitivity and specificity for differentiating between cases of metopic synostosis and controls, several measures had comparable diagnostic capacity with AUC values above 0.9. The scaled MSS had an AUC of 0.96 ($P < 0.001$;

Table 2. Multivariable Linear Regression Model for Scaled MSS (R-Square = 0.66).

Variable	P	Standardized B
Adjusted EBA	0.023	−0.16
Frontal coronal angle	0.64	−0.36
Posterior angle of the anterior fossa	<0.001	0.26
Orbital rim angle	0.15	0.10
Temporal depression angle	0.042	−0.12
Age (months)	<0.001	−0.37
Biparietal distance	<0.001	0.57
Interdacryon distance	0.033	−0.15
Orbital width	<0.001	−0.31
Male	0.41	0.05

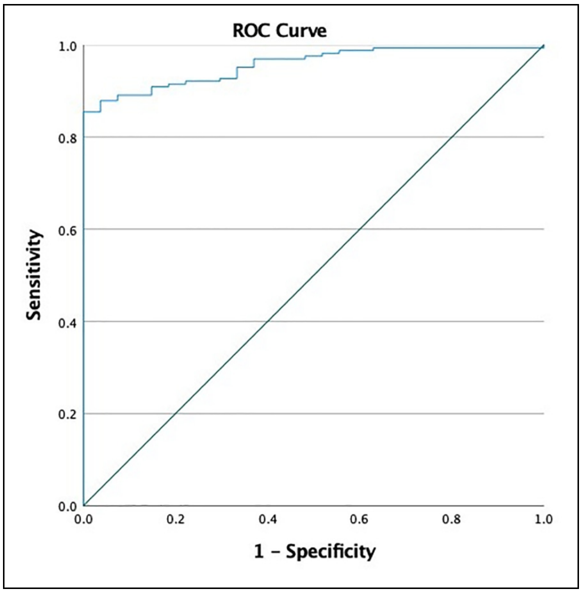

Figure 3. Receiver operating characteristic curve displaying the diagnostic value of the machine learning algorithm generated metopic severity score for distinguishing metopic craniosynostosis from controls.

Figure 3). The value for scaled MSS that maximized sensitivity and specificity was 2.16, which resulted in a sensitivity of 0.891 and a specificity of 0.926.

The strongest individual measurements were the adjusted EBA, EBA, and biparietal/bitemporal ratio which had AUC values of 0.98 ($P < 0.001$), 0.97 ($P < 0.001$), and 0.95 ($P < 0.001$). The other craniometric indices were less valuable. The frontal coronal angle ($AUC = 0.798$, $P < 0.001$), upper orbital width ($AUC = 0.75$, $P < 0.001$), and interdacryon distance ($AUC = 0.74$, $P < 0.001$) had lower but still statistically significant AUC values (Table 3).

Table 3. ROC Analysis for Each Measurement in Distinguishing Cases from Controls.

Variable	Area	Standard Error	P
Adjusted EBA	0.98	0.008	<0.001
EBA	0.974	0.011	<0.001
ScaledMSS	0.957	0.014	<0.001
Bitemporal/Biparietal Ratio	0.935	0.023	<0.001
Frontal Coronal Angle	0.798	0.044	<0.001
Upper Orbital Width	0.752	0.043	<0.001
Interdacryon Distance	0.742	0.043	<0.001
Interzygomaticofrontal suture distance	0.66	0.07	0.008
Lateral Orbital Width	0.591	0.071	0.131
Posterior Angle of Anterior Fossa	0.556	0.06	0.35
AP Diameter	0.547	0.076	0.438
Metopic Index	0.532	0.059	0.589
Horizontal Cone Angle	0.506	0.057	0.921
Foramen Ovale Distance	0.476	0.071	0.691
Orbital Rim Angle	0.462	0.057	0.529

Discussion

The operative indication for metopic synostosis remains undefined as the metopic suture is the only major suture to close physiologically in infancy. Various attempts to define a threshold for intervention based on degree of morphologic deformity have been published. Recently, machine learning technology has been combined with 3-D imaging data to provide an automated measure of severity not limited by 2-D manipulation of the image or human measurement error.

This study sought to explore which manual measurements most correlated with algorithm severity, as well as determine the efficacy of each metric in distinguishing scans of metopic synostosis from control scans. Posnick et al., first quantified various measurements that significantly differed in metopic synostosis on craniomorphometric analysis including bitemporal width, interorbital distance, and medial wall protrusion. (Posnick et al., 1994) Subsequent studies proposed various severity metrics based on manual measurements of CT imaging, including the EBA proposed by Beckett et al., (Beckett et al., 2012) adjusted EBA by Naran et al., (Naran et al., 2017) metopic index by Wang et al., (Wang et al., 2016) orbital rim angle by Ezaldein et al., (Ezaldein et al., 2014) intercranial volume by Cronin et al., (Cronin et al., 2020) central angle by Havlik et al., (Havlik et al., 1999) and frontal angle by Chandler et al.(Chandler et al., 2021).

The high R-square of 0.66 of the linear regression model developed in this study indicated a high degree of correlation between the composite of the manual measurements and the scaled MSS. **While significant individual measurements exhibited a moderate correlation with scaled MSS (absolute value of R between 0.3–0.5),** the multivariable model demonstrated a strong correlation with high predictive capacity. Thus, the algorithm provided a holistic evaluation of severity that considered several of the previously published manual metrics.

One previous study has attempted to characterize severity in metopic synostosis by considering global head shape. Ruiz-Correa et al., proposed a geometrically-based approach for characterizing the deformity in metopic synostosis at multiple levels rather than relying on an individual plane.(Ruiz-Correa et al., 2008) The Trigonocephaly Severity Index was based on outlines of the calvaria calculated at various axial planes, and thus incorporated overall information regarding the skull shape. The composite of manual measurements computed in TSI proved to be superior individual measurements based on an isolated measurement on a single slice of CT. The drawback to the method proposed by Ruiz-Correa and colleagues, however, was the complex nature of the calculation which relied on software not readily available at all institutions. The scaled MSS computer-generated score is similarly based on skull shape at a global level, and thus would provide an attractive alternative for assessing severity that is much easier to incorporate into practice.

Of the manual metrics included in this study, the algorithm score correlated most with biparietal distance, orbital width, and age. The positive correlation between biparietal distance and severity score likely represents the compensatory increase in lateral diameter that result with more severe constriction of the frontal bone. The fact that age negatively correlated with scaled MSS also suggests a relationship between age at presentation and severity of deformity. Study of the natural history of metopic suture synostosis has shown that suture fusion begins at about 4 months in the majority of cases with almost ubiquitous closure after 9 months.(Pindrik et al., 2016; Vinchon, 2019) Fusion typically commences at the nasion and proceeds superiorly.(Weinzweig et al., 2003) This study suggests a relationship between age of suture fusion and severity of morphologic deformity, as more morphologically severe patients had an earlier presentation.

Many of the indices performed well in distinguishing cases from controls, including scaled MSS. Of the measurements tested, adjusted EBA, EBA, biparietal/bitemporal ratio, and scaled MSS all were comparable in differentiating the cases. The advantages of using the computer algorithm are the reduction of the risk of measurement error as well as the ease of application. The quantitative methods to measure severity manually rely on the calculation of distances and angles based on selected anatomic landmarks which may be difficult to locate or can be affected by image orientation. **The measurements in this study were analyzed only after orienting each skull in a consistent plane using specialized software, which may not be readily available.** While no algorithm can replace physician judgement in individual cases, the algorithm offers an objective means by which physicians can assess morphologic deformity and can serve as an adjunct in clinical decision-making for ambiguous cases.

The cases of metopic craniosynostosis included in this study were compiled from multiple US craniofacial centers and were operative cases of metopic craniosynostosis. One limitation of this study is that the diagnosis of metopic synostosis was dependent on the judgement of the surgeon who contributed the case to the database. In addition, this study included only patients who underwent surgery for metopic craniosynostosis. Future studies seek to include patients who were labeled as “mild” cases and in whom surgery was deferred. Most mild cases of metopic synostosis who do not undergo surgery usually will not undergo imaging. This study nevertheless presents the largest cohort of patients with metopic synostosis evaluated by a machine learning algorithm.

Conclusion

A machine learning algorithm for determining severity in metopic synostosis has the advantage over previously described craniometric measures by providing a composite evaluation of skull shape abnormality. The severity score generated by the machine learning algorithm is comparable to previously validated severity indices in distinguishing cases of metopic synostosis from controls.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

ShapeWorks was supported by the National Institutes of Health under grant numbers NIBIB-U24EB029011, NIBIB-R01EB016701, NIAMS-R01AR076120, NHLBIR01HL135568, and NIGMS-P41GM103545 and R24 GM136986.

ORCID iDs

Madeleine K. Bruce https://orcid.org/0000-0002-4021-4551
Jesse A. Goldstein https://orcid.org/0000-0002-5242-3844
Michael Alperovich https://orcid.org/0000-0002-5093-5790

Supplemental Material

Supplemental material for this article is available online.

References

Beckett JS, Chadha P, Persing JA, Steinbacher DM. Classification of trigonocephaly in metopic synostosis. *Plast Reconstr Surg*. 2012;130(3):442e-447e.

Bhalodia R, Dvoracek LA, Ayyash AM, Kavan L, Whitaker R, Goldstein JA. Quantifying the severity of metopic craniosynostosis: a pilot study application of machine learning in craniofacial surgery. *J Craniofac Surg*. 2020;31(3):697-701.

Birgfeld CB, Saltzman BS, Hing AV, Heike CL, Khanna PC, Gruss JS, Hopper RA. Making the diagnosis: metopic ridge versus metopic craniosynostosis. *J Craniofac Surg*. 2013;24(1):178-185.

Bottero L, Lajeunie E, Arnaud E, Marchac D, Renier D. Functional outcome after surgery for trigonocephaly. *Plast Reconstr Surg*. 1998;102(4):952-958; discussion 959–960.

Cates J, Elhabian S, Whitaker R. Shapeworks: particle-based shape correspondence and visualization software. In: *Statistical Shape and Deformation Analysis*. Academic Press, 2017; 257–298. https://www.sci.utah.edu/cibc-software/shapeworks.html

Chandler L, Park KE, Allam O, Mozaffari MA, Khetpal S, Smetona J, Pourtaheri N, Lu X, Persing JA, Alperovich M. Distinguishing craniomorphometric characteristics and severity in metopic synostosis patients. *Int J Oral Maxillofac Surg*. 2021;50(8):1040–1046.

Cronin BJ, Brandel MG, McKee RM, Hashmi A, Oviedo P, Buckstaff T, Cahill G, Mannix E, Reid CM, Lance S, et al. A comparison of intracranial volume growth in normal children and patients With metopic craniosynostosis. *J Craniofac Surg*. 2020;31(1):142-146.

Ezaldein HH, Metzler P, Persing JA, Steinbacher DM. Three-dimensional orbital dysmorphology in metopic synostosis. *J Plast Reconstr Aesthet Surg*. 2014;67(7):900-905.

Havlik RJ, Azurin DJ, Bartlett SP, Whitaker LA. Analysis and treatment of severe trigonocephaly. *Plast Reconstr Surg*. 1999;103(2):381-390.

Kellogg R, Allori AC, Rogers GF, Marcus JR. Interfrontal angle for characterization of trigonocephaly: part 1: development and validation of a tool for diagnosis of metopic synostosis. *J Craniofac Surg*. 2012;23(3):799-804.

Kolar JC. An epidemiological study of nonsyndromal craniosynostoses. *J Craniofac Surg*. 2011;22(1):47-49.

Naran S, Mazzaferro D, Wes A, Vossough A, Bartlett SP, Taylor JA. A craniometric analysis of cranial base and cranial vault differences in patients With metopic craniosynostosis. *J Craniofac Surg*. 2017;28(8):2030-2035.

Pindrik J, Molenda J, Uribe-Cardenas R, Dorafshar AH, Ahn ES. Normative ranges of anthropometric cranial indices and metopic suture closure during infancy. *J Neurosurg Pediatr*. 2016; 25(6):667-673.

Posnick JC, Lin KY, Chen P, Armstrong D. Metopic synostosis: quantitative assessment of presenting deformity and surgical results based on CT scans. *Plast Reconstr Surg*. 1994;93(1):16-24.

Ruiz-Correa S, Starr JR, Lin HJ, Kapp-Simon KA, Sze RW, Ellenbogen RG, Speltz ML, Cunningham ML. New severity indices for quantifying single-suture metopic craniosynostosis. *Neurosurgery*. 2008;63(2):318-324; discussion 324–315.

Vinchon M. The metopic suture: natural history. *Neurochirurgie*. 2019;65(5):239-245.

Vu HL, Panchal J, Parker EE, Levine NS, Francel P. The timing of physiologic closure of the metopic suture: a review of 159 patients using reconstructed 3D CT scans of the craniofacial region. *J Craniofac Surg*. 2001;12(6):527-532.

Wang JY, Dorafshar AH, Liu A, Groves ML, Ahn ES. The metopic index: an anthropometric index for the quantitative assessment of trigonocephaly from metopic synostosis. *J Neurosurg Pediatr*. 2016;18(3):275-280.

Weinzweig J, Kirschner RE, Farley A, Reiss P, Hunter J, Whitaker LA, Bartlett SP. Metopic synostosis: defining the temporal sequence of normal suture fusion and differentiating it from synostosis on the basis of computed tomography images. *Plast Reconstr Surg*. 2003;112(5):1211-1218.