As many as 1 in 2 children in North America will develop a cranial deformity (i.e., misshapen head) by 2 months of age. In most cases, these cosmetic deformities will resolve with observation or physical therapy by the time the child turns 5 years of age, but severe cases require the addition of orthotic therapy (fitted helmets) and, in the case of patients with early fusion of the bones in their skulls (i.e., craniosynostosis), surgery to release the fused bones.1,2

Therapeutic interventions for patients with cranial deformities are time critical. The efficacy of orthotic and surgical options dwindles as the child gets older. The most effective and least invasive therapies are best performed before the child turns 4–6 months of age.3 After the child is 12 months of age, most therapeutic interventions are no longer effective.3 Early detection and referral to specialists are therefore key to improving the management of cranial shape pathologies.

While deformational cranial deformity shapes are quite familiar to pediatricians, the head shapes most in need of review by a pediatric craniofacial specialist—craniosynostosis shapes—are frequently not identified until the child is too old to be a candidate for the least-invasive surgical treatments.4 For those cranial shapes least familiar to primary care providers, there are no in-office tests capable of establishing the diagnosis.

Until now, identification of a pathologic head shape had to be made by a qualified specialist who visually inspected
the child’s head or reviewed features of a high-quality radiographic adjunct. The analysis has been entirely driven by human pattern recognition.

Modern image processing techniques now allow for the systematic analysis and comparison of images—essentially replicating human visual pattern recognition. Leveraging these digital techniques, software can be developed to parse, measure, and interpret single- or multiple-perspective images of heads of pediatric patients. Outputs from this software can then provide pediatricians, family medicine doctors, craniofacial surgeons, and parents with metrics for cranial shape severity that can be used to assess the probability that the patient’s head fits within a pathologic category requiring referral to a specialist. This software could easily be incorporated into phone-, desktop-, or web-based platforms, making it easy to transmit images to end users and reducing dependence on craniofacial providers and radiological adjuncts for screening.

We present preliminary results from craniometric software designed to acquire conventional and novel craniometrics from digital patient images and parse them through machine learning classification algorithms. The diagnostic accuracy of a common machine learning model, constructed from software-derived craniometrics, compared to craniofacial specialist assessment is reported, as are intraclass correlation coefficients for component craniometrics compared to current standard craniometric equipment.

**Methods**

**Image Selection**

A Google search was performed using the keywords “craniosynostosis,” “plagiocephaly,” “brachycephaly,” and “normal newborn head.” Forty open-source digital images were selected, with at least one image corresponding to each of the groups listed in Fig. 1. Only orthogonal, top-down images and face images without disruption of the margins of the head’s equator were admitted for this study. These images were used for initial software development and model training, though not all samples included for development included both image types.

A set of additional retrospective patient images, conforming to the same format listed above, were culled from the craniofacial and neurosurgery clinic medical records of patients presenting to Connecticut Children’s, Mayo Hospital, or Johns Hopkins Hospital for evaluation of “abnormal head shape,” “craniosynostosis,” or “plagiocephaly” between April 2008 and February 2020 (IRB approval no. 19-047). Patients < 1 year of age at the time of the image capture were eligible for inclusion. With regard to top-down photos, significant obscuration of the equator of the head by a solid object or tilting of the head such that the immediate margins of the euryon, glabella, or opisthocranion were not identifiable constituted exclusion criteria. Facial photos were excluded if the Haar cascade could not identify a “face” within the image. Additionally, optical scan craniometric data, age in months at the time of encounter, and sex were collected from each patient record.

Images were classified by size (pixels x pixels), dots per inch, type of head shape (adjudicated by the authors), and whether the child grossly had < 75% or > 75% scalp coverage by hair.

**Craniometric Acquisition**

All image processing and software-based craniometric measurements were performed using a C# library designed by the lead author (M.J.B.) in Visual Studio 2017 (Microsoft).

Foreground extraction of the heads within the image was achieved using a function based on the GrabCut model. Face identification and automated landmark extraction were achieved with a Haar cascade and utilization of portions of the EMGU image processing library.

**Craniometric Measurements**

The cephalic index (CI) and cranial vault asymmetry index (CVAI) were calculated in the standard fashion (Fig. 2A and B, Table 1). Optical scanner–derived CI and CVAI measurements were pulled from the medical record if they were acquired within 7 days of a top-down cranial image for the same subject.

Hu moments for the 7 cranial head types examined (Fig. 1) were calculated (eq. 1).7

\[
m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x,y) \, dx \, dy, \quad q, p = 0, 1, 2, \ldots \quad [\text{eq. 1}]\]

The contours of the most severe examples of each of the 7 cranial head types in the open-source data set, as determined by the lead author (M.J.B.), were utilized in generating the template Hu moments for each shape. Comparisons were calculated using the euclidean distance between the invariant (1–6) log-transformed Hu moments (H) of the two shapes/contours being evaluated (eq. 2).

\[
D(A,B) = \frac{\sqrt{\sum_{i=0}^{6} (H_i^A - H_i^B)^2}}{\sqrt{6}} \quad [\text{eq. 2}]\]

Measurements and calculations for novel craniometrics evaluated in this study are presented in Fig. 2 and Table 1.

**Head Shape Classification**

Open-source cranial images were classified by the attribution given them from their source, if the lead author concurred. Prior to craniometric analysis, retrospective patient images processed from clinical cases were reviewed in a blinded fashion by the authors (M.J.B., J.E.M., E.S.A.) and classified according to the major head shape types listed in Fig. 1. Discordant classifications were re-evaluated by the lead author (M.J.B.), and if disagreement remained, the case was discarded from analysis.

**Statistical Analysis**

Scalar paired data were evaluated with a Student t-test. Interrater differences in CI and CVAI measurements generated by optical scans compared to the craniometric
software were evaluated with intraclass correlation coefficients. Craniometrics for individual head shape types were analyzed using receiver operating characteristic (ROC) analysis. All statistics were calculated using the IBM SPSS software version 26 (IBM Corp.) and RStudio version 1.2.

Classification modeling was performed using linear discriminant analysis. A fivefold cross-validation was used to assess model performance within a subset of 60 subjects with complete top-down and facial photo data points, including all cranial classification types listed in Fig. 1. Models were derived from RStudio using the “MASS” package.

**Results**

**Demographics**

The cranial shape classification software accuracy was compared to that of blinded craniofacial specialists using a data set of open-source (n = 40) and retrospectively collected newborn orthogonal top-down cranial images, with or without additional facial views (n = 339). There was a sex distribution of 2.4 males to 1 female within the retrospective clinical data set of images. The mean ± SD age at the time clinical photos were acquired was 136.7 ± 78.2 days.

The frequency of different cranial head shape types based on expert reviewer classification is described in Fig. 3A. Given the paucity of bicoronal and lambdoid craniosynostosis cases identified in the initial pool of images, these classes were excluded from analysis. Five cases were excluded from further analysis due to a lack of reviewer agreement. Two cases carried diagnoses of metopic craniosynostosis, and 3 cases carried diagnoses of plagiocephaly within their respective medical records. There were no discrepancies between reviewer craniosynostosis head shape classifications and craniosynostosis diagnoses noted within the medical records of the training or test data sets.
Software-Derived Conventional Craniometrics Versus Industry Standards

Software-generated CI and CVAI measurements were compared to optical scanner-derived measurements. The optical scanner CI and CVAI measurements had excellent ($\kappa = 0.95$, 95% CI 0.84–0.98; $p < 0.001$) and good ($\kappa = 0.67$, 95% CI 0.24–0.88; $p = 0.003$) levels of agreement with the software-generated CI and CVAI measurements, respectively (eFig. 1).

Craniometrics of Conventional Head Shape Types

ROC analysis of each craniometric captured by the software was performed for the most frequent cranial head shape types (eTable 1). Normal-type head shapes were not well defined by any of the craniometrics examined. CI had excellent discriminatory power for brachycephaly subjects, with a CI > 93.7% being 92.5% sensitive and 89.9% specific. The next most effective and significant craniometrics were the brachycephaly Hu moment euclidean distance using a cutoff < 1.44 and the posterior arc angle (PAA) where values were > 2.71 rad. Plagiocephaly in this cohort was well classified by CVAI > 5.63, with 76.7% sensitivity and 75.9% specificity. Sagittal craniosynostosis head shape types had multiple highly significant cranio-

metrics ($p < 0.001$), including a CI < 75.3%, a sagittal Hu moment euclidean distance < 0.20, a brachycephaly Hu moment euclidean distance > 2.13, and a PAA < 2.40 rad. For trigonocephaly subjects, an anterior arc angle (AAA) < 2.28 rad was highly (> 90%) sensitive and specific, while an anterior-middle width ratio (AMWR) < 0.82 provided a less robust sensitivity of 77.9% and specificity of 83.5%. Unicoronal head shape types were significantly ($p < 0.02$) associated with transcanthal line angle (TCLA). TCLAs > 1.61 rad or < 1.53 rad, respectively, had 100.0% sensitivi-
ties and > 95.0% specificities for left and right unicoronal craniosynostoses.

Diagnostic Model for Head Shape Classification Using Machine Learning

Component variables within the linear discriminant model were pruned based on a variance inflation factor < 5 (eTable 2). Cross-validated accuracy for the linear discriminant model was 88.3% (95% CI 77.4%–95.2%; $p < 0.001$). Terminal node classification accuracies are detailed in Table 2. The cumulative accuracy for craniosynostosis deformities within the model was 93.3% (95% CI 86.8–99.8; $p < 0.001$) with a sensitivity of 92.0% and specificity of 94.3% (Fig. 3B).

FIG. 2. Illustration of craniometric measurement acquisition. A: Schematics of the measurements used in calculating top-down craniometrics, including CI and CVAI. Novel top-down craniometrics for the left, right, and total anterior and posterior arc angles, as well as the anterior-middle, posterior-middle, and anterior-posterior width ratios. B: The facial craniometric TCLA is diagrammed. Ant = anterior; AP = anterior-posterior; LH = left hemisphere; Mid = middle; ML = midline; Post = posterior; RH = right hemisphere. Figure is available in color online only.
TABLE 1. Craniometric calculations

<table>
<thead>
<tr>
<th>Angular and Ratio-Based Craniometrics</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional craniometrics</td>
<td></td>
</tr>
<tr>
<td>CI</td>
<td>( CI = \frac{ML}{AP} \times 100 )</td>
</tr>
<tr>
<td>CVAI</td>
<td>( CVAI = \frac{</td>
</tr>
<tr>
<td>Novel craniometrics</td>
<td></td>
</tr>
<tr>
<td>AMWR</td>
<td>( AMWR = \frac{Ant}{Mid} )</td>
</tr>
<tr>
<td>APWR</td>
<td>( APWR = \frac{Ant}{Post} )</td>
</tr>
<tr>
<td>LRHR</td>
<td>Numerator ((N) = \text{lt AP length (1/4 diameter lt of ML)} ) or ( \text{rt AP length (1/4 diameter rt of ML)} ), whichever is smaller</td>
</tr>
<tr>
<td></td>
<td>Denominator ((D) = \text{lt AP length or rt AP length, whichever is larger} )</td>
</tr>
<tr>
<td></td>
<td>( LRHR = \frac{N}{D} )</td>
</tr>
</tbody>
</table>

A = line A; Ant = anterior; AP = anterior-posterior; APWR = anterior-posterior width ratio; B = line B; LRHR = left-right height ratio; Mid = middle; ML = midline; Post = posterior.

Conventional and novel angular and ratio-based craniometrics employed for cranial deformation evaluation and classification using the proposed craniometric software. Measurement locations are shown in Fig. 2.

Discussion

Objective, easily acquired, and reproducible craniometrics for the gamut of clinically relevant pediatric cranial shapes are essential to the development of reliable screening and outcome analyses for congenital cranial deformities. Unfortunately, widely available craniometrics for the majority of craniosynostosis and positional deformities are still lacking, with CVAI and CI being the only measurements in use that can practically be acquired in the ambulatory setting. Some groups have begun to leverage measurements in use that can practically be acquired in the ambulatory setting.8,9 Some groups have begun to leverage measurements in use that can practically be acquired in the ambulatory setting.8,9

Software Fidelity

All craniometrics examined in this study are based on vectors, angles, or ratios of absolute measurements, so they can be compared across observations with little regard for the observer’s distance relative to the subject or the subject’s absolute size.

This is not to say that observer perspective does not impact craniometric calculations. Image orientation is critical to the ratio-based craniometrics evaluated here. We found that disturbances in image orientation around the cranio-caudal axis of the patient of more than 5° off midline lead to marked variability in the ratio-based cranial measurements (eTable 3). However, data from test-user performance trials during software development indicate that users produce < 5° of rotation error > 95% of the time (eFig. 2). This suggests that, in a real-world application of this software, rotation error is unlikely to impact software output.

Extracted head shape area fidelity was also robust, with exceedingly high intra-user agreement on sample head contours (eFig. 3). It should be cautioned, though, that intra-user agreement does not imply accuracy, and increasing disruption of the head circumference contour by hair or clothing would reasonably be expected to distort the contours acquired by this software. Such corruptions of the extracted head shape contours will, naturally, make the craniometry and head shape classification outputs unreliable. Similarly, facial craniometrics rely on key landmarks being distinguishable within the camera’s field of view, and we encountered a small subset of images within our data set that had to be excluded because of an abundance of obscuration within the frame.

Finally, with regard to software craniometric accuracy, it is very encouraging that the software was able to produce very good levels of agreement with CI and CVAI measurements acquired through optical scans. Optical scanners are the measurement tools of choice for most orthotics groups and are often the source of craniometrics reported to insurance companies and those used to inform the response to cranial orthosis therapy or surgery. Given the high degree of interrater agreement noted between this software and the gold standard craniometry diagnostic device, we feel encouraged that this software can provide accurate and meaningful diagnostic feedback for users. If incorporated into a smartphone application, this tool could allow families and providers questioning the severity of a child’s cranial deformity to simply acquire a set of orthogonally oriented photos of the child’s head and face, and the software could then provide CI (scaphocephaly, brachycephaly), CVAI (plagiocephaly), AAA (trigonocephaly), and TCLA (anterior congenital plagiocephaly) measurements that can be trended over time and interpreted for the user in terms of normative distributions.

Individual Craniometric Accuracy

To explore the potential for utilizing these novel cranio-
FIG. 3. Head shape frequencies and sample flow through index and reference examinations. 

A: Frequency of head shape types within the total data set, based on expert classification. B: Flow of cranial images through index test development and performance with a cross-validated testing set.

Unauthenticated | Downloaded 12/02/23 10:32 PM UTC
metrics to identify cranial deformities via computationally driven algorithms, we calculated ROC analyses for each craniometric with respect to “normal,” scaphocephalic, brachycephalic, plagiocephalic, unicoronal, and trigonocephalic head shapes. Our results confirmed the value of CI for the classification of scaphocephaly and brachycephaly, with our normative range of 75%–93% aligning closely with previously described normal CI ranges of 75%–95%. The widely described CVAI was also highly effective in discriminating plagiocephaly patients within this study, with no comparable craniometric testing in this study rising to its level of sensitivity and specificity.

More interesting, though, are the novel contour-based craniometrics utilized in this study. With scaphocephaly subjects, PAA and sagittal Hu moment euclidean distance provided comparable levels of exceptional sensitivity and specificity to CI, while brachycephaly was effectively screened using the brachycephaly Hu moment euclidean distance and PAA. The accuracy of Hu moment–based craniometrics in this study creates the potential for tailoring cranial image processing–based software to orientation-insensitive metrics that can operate on a wider range of subject orientations.

Perhaps the most interesting cranial deformity groups that we reviewed in this pilot study were infants with trigonocephaly and congenital anterior plagiocephaly, which are cranial deformities that typically require surgery but do not have standard craniometrics that can be evaluated without high-resolution 3D CT. We identified AAA and TCLA as significant discriminators between trigonocephaly images and anterior plagiocephaly images, respectively, each with > 90% sensitivities and specificities. The advantage of having access to these highly discriminatory measurements in a rapidly deployable phone or desktop-based software application cannot be overstated, given that such measurements have previously been accomplished through expensive CT scans acquired at sites located at a distance from patient homes and primary care clinics.

Application of Classification Algorithms for Categorizing Head Shapes

Our approach to operationalizing the above craniometrics for the diagnosis of newborn cranial deformities was to trial a common machine learning algorithm—linear discriminant analysis. While this method may not prove to be the most effective model in the long term, the size of the current data set, number of variables, and discriminatory strength of our craniometrics recommended this approach as a reasonable starting point compared to more robust systems, such as random forest.

Using linear discriminant modeling, this diagnostic software was able to discern normal or deformational head shapes from potentially surgical craniosynostosis head shapes with 93.3% accuracy. Even when using this model to focus on specific head shape types, the software was able to accurately identify cranial shapes 88.3% of the time. These very encouraging results can certainly be improved upon with a larger, more representative training data set. However, even as the model stands now, this digital image–based method for discriminating craniosynostosis head shapes versus noncraniosynostosis head shapes, using the linear discriminant method, could offer significant value to parents and pediatricians seeking a cheap and broadly deployable screening tool. The authors believe that recursive model generation with an expanded training data set is the next step in refining real-time image-based craniometrics for the screening and evaluation of congenital cranial deformities.

Study Limitations

While linear discriminant analysis proved effective at discerning synostotic from nonsynostotic head shapes, as well as correctly identifying individual head shape types, the data set analyzed lacked less common subtypes of craniosynostosis, to include bicoronal, lambdoid, and polysutural craniosynostoses. As a result, this diagnostic tool cannot be generalized to the roughly 5% of nonsyndromic cases falling into the aforementioned categories.

Additionally, the reliance on open-source and referral center patient images means that this training data set likely does not completely represent the relative distribution of head shape types found in primary care settings or certainly in the population at large.

Conclusions

Machine learning algorithms for evaluating head shapes of newborns can produce diagnostic accuracies nearly on par with those of specialist evaluations. Additionally, exceptionally reliable craniometrics are achievable through image processing techniques, allowing for point-of-care measurements for orthotic referrals and assessment of cranial deformity treatment responses. These pilot data suggest that future iterations of this digital approach to newborn cranial deformity screening could provide remote screening capabilities for primary care practices engaged in the telehealth arena, bringing craniometry capabilities to providers and families, and, possibly, reduce radiation exposures and diagnostic costs.
Acknowledgments

The authors thank David Hersh, MD, Eva Johnson, and Allison Labelle, APRN, for their assistance and feedback during the software development, including performance testing and refinement of the user interface.

References


Disclosures

The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author Contributions

Conception and design: Bookland. Acquisition of data: Ahn, Stoltz. Analysis and interpretation of data: Bookland. Drafting the article: Bookland, Martin. Critically revising the article: Ahn, Martin. Reviewed submitted version of manuscript: Bookland, Ahn, Martin. Approved the final version of the manuscript on behalf of all authors: Bookland. Statistical analysis: Bookland. Administrative/technical/material support: Stoltz. Study supervision: Bookland.

Supplemental Information

Online-Only Content

Supplemental material is available with the online version of the article.


Correspondence

Markus J. Bookland: Connecticut Children’s, Hartford, CT. mbookland@connecticutchildrens.org.