

# New Scaphocephaly Severity Indices of Sagittal Craniosynostosis: A Comparative Study With Cranial Index Quantifications

Salvador Ruiz-Correa, Ph.D., Raymond W. Sze, M.D., Jacqueline R. Starr, Ph.D., Hen-Tzu J. Lin, M.S., Matthew L. Speltz, Ph.D., Michael L. Cunningham, M.D., Ph.D., Anne V. Hing, M.D.

**Objective:** To describe a novel set of scaphocephaly severity indices (SSIs) for predicting and quantifying head- and skull-shape deformity in children diagnosed with isolated sagittal synostosis (ISS) and compare their sensitivity and specificity with those of the traditional cranial index (CI).

**Methods:** Computed tomography head scans were obtained from 60 patients diagnosed with ISS and 41 age-matched control patients. Volumetric reformations of the skull and overlying skin were used to trace two-dimensional planes defined in terms of skull-base plane and internal or surface landmarks. For each patient, novel SSIs were computed as the ratio of head width and length as measured on each of these planes. A traditional CI was also calculated and a receiver operating characteristic curve analysis was applied to compare the sensitivity and specificity of the proposed indices with those of CI.

**Results:** Although the CI is a sensitive measure of scaphocephaly, it is not specific and therefore not a suitable predictor of ISS in many practical applications. The SSI-A provides a specificity of 95% at a sensitivity level of 98%, in contrast with the 68% of CI. On average, the sensitivity and specificity of all proposed indices are superior to those of CI.

**Conclusions:** Measurements of cranial width and length derived from planes that are defined in terms of internal or surface landmarks and skull-base plane produce SSIs that outperform traditional CI measurements.

KEY WORDS: *cephalic index, cranial base, cranial index, craniosynostosis, isolated sagittal synostosis, scaphocephaly severity index, shape analysis, skull-base plane*

Sagittal synostosis is the most common form of isolated suture synostosis, with an incidence of approximately 1 in 5000 and accounting for 40% to 60% of single suture synostoses (Hunter and Rudd, 1976; Lajeune et al., 1996). Premature closure of the sagittal suture results in scaphocephaly, de-

noting a long narrow skull with more-or-less prominent ridges along the prematurely ossified sagittal suture. The degree of suture fusion, severity of head-shape deformity, and additional changes such as frontal and occipital bossing and biparietal narrowing can vary significantly among affected individuals (Fig. 1).

The degree of severity of isolated sagittal synostosis (ISS) is commonly quantified by measuring the cranial index (CI). The CI, first used in 1842 by the Swedish anatomist Andreas Retzius (Kolar et al., 1997), represents the ratio of maximum cranial width to maximum cranial length. Children with ISS typically have an average CI of 60% to 67% (David and Simpson, 1982; Kaiser, 1988; Slomic et al., 1992; Farkas et al., 1994; Solan et al., 1997; Panchal et al., 1999; Christophis et al., 2001; Guimaraes-Ferreira et al., 2001), whereas children with normal head shape have an average CI of 76% to 78% (Slomic et al., 1992; Farkas et al., 1994). Intracranial volume is also a measure of ISS severity that has been used in clinical research and can be directly derived from two-dimensional (2D) computed tomography (CT) measurements (Posnick et al., 1992, 1993, 1995; Waitzman et al., 1992).

The diagnosis of ISS is typically made on the basis of clinical judgment, with imaging to confirm the clinician's impres-

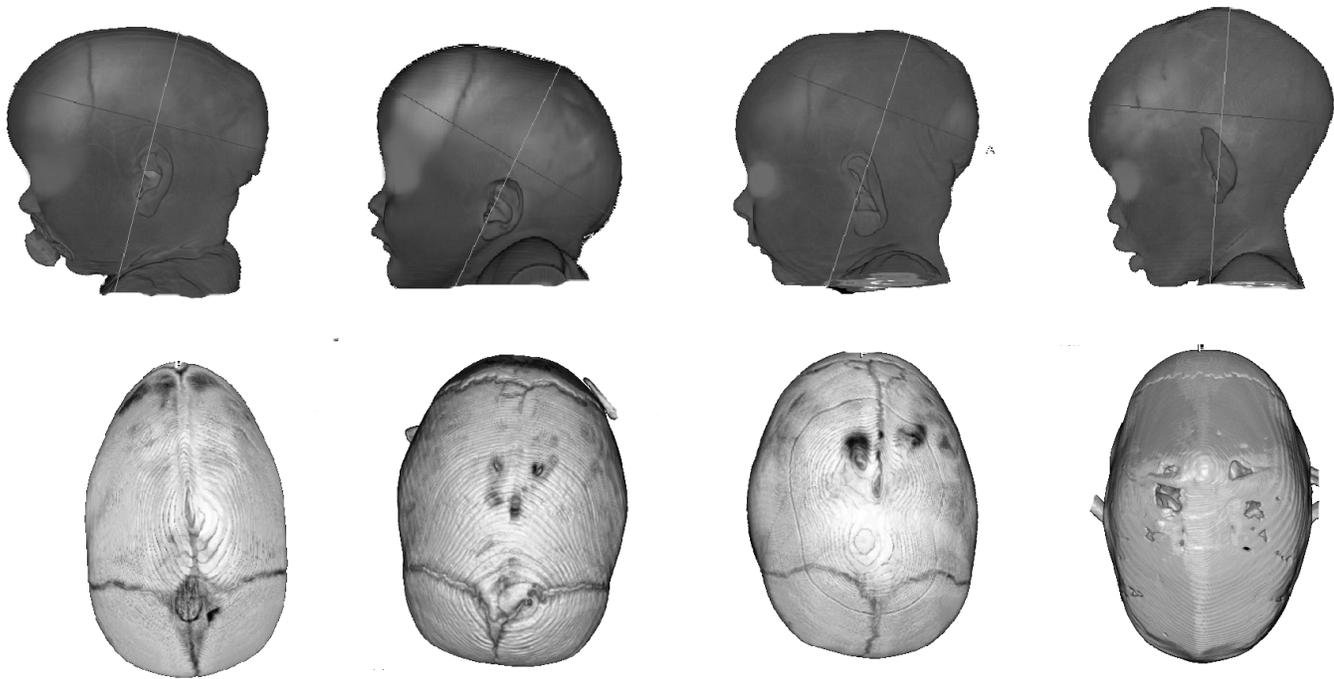
---

Drs. Ruiz-Correa and Sze are with the Department of Radiology, University of Washington, Seattle, Washington, and the Department of Radiology, Children's Hospital and Regional Medical Center, Seattle, Washington. Drs. Starr and Cunningham are with the Children's Craniofacial Center, Children's Hospital and Regional Medical Center, Seattle, Washington, and the Department of Pediatrics and Epidemiology, University of Washington, Seattle, Washington. Ms. Lin is with the Department of Biomedical and Health Informatics and Dr. Speltz is with the Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, Washington. Dr. Hing is with the Children's Craniofacial Center, Children's Hospital and Regional Medical Center, Seattle, Washington.

This work was supported in part by The Laurel Foundation Center for Craniofacial Research, The Marsha Solan-Glazer Craniofacial Endowment, and a grant from the National Institute of Dental and Craniofacial Research (NIDCR grant R01 DE 13813 awarded to Dr. Matthew Speltz).

Submitted January 2005; Accepted July 2005.

Address correspondence to: Salvador Ruiz-Correa, Children's Hospital and Regional Medical Center, Department of Radiology R54-38, 4800 Sand Point Way NE, Seattle, WA 98105. E-mail [sruiz@u.washington.edu](mailto:sruiz@u.washington.edu).



**FIGURE 1** Three-dimensional reformations of the head of patients affected with ISS. Lateral view of overlying skin (top row). Top view of the skull (bottom row).

sion. Although quantitative indices of head shape such as CI or intracranial volume are not often used for clinical diagnosis, they have been used by researchers to compare the timing (Panchal et al., 1999) and outcome of different surgical procedures (Kaiser et al., 1988; Marsh et al., 1991; Posnick et al., 1992, 1993, 1995; Waitzman et al., 1992; Panchal et al., 1999; Fata and Turner, 2001; Guimaraes-Ferreira et al. 2001, 2003) and sometimes in combination with cranial molding (Seymour-Dempsey et al., 2002; David et al., 2004; Jimenez et al., 2004). A variety of methods for obtaining the CI have been reported in the surgical literature. The most common method involves using digital calipers and maximum cranial width and length from skull radiographs (including cephalograms, xerograms, or three-dimensional [3D] CT images).

Variation in the methodology used to determine cranial width and length has limited comparisons among studies. For example, to quantify maximum cranial width, the caliper technique requires the identification of the euryon (the most lateral point on each side of the head). Similarly, the determination of the cranial length requires the somewhat arbitrary identification of the glabella, the most prominent point in the medial sagittal plane between the supraorbital ridges, and the opisthocranium, the most prominent point of the occiput (see Kolar et al., 1997; Cohen and MacLean, 2000). Determination of cranial length using calipers is also complicated by scaphocephalic head shape itself, as bossing of the forehead region superior to the glabella may distort the glabella-opisthocranium axis so that it no longer represents the maximum cranial length (Albright et al., 1996; Friede et al., 1996; Guimaraes-Ferreira et al., 2001, 2003). Identification of specific anterior landmarks (glabella) has not been used in radiographic mea-

asures of maximum cranial length, and when CI has been obtained from 3D CT images, there has been variation associated with both the specific view taken (i.e., frontal or lateral view for cranial width measures; Panchal et al., 1999; Fata and Turner, 2001) and the type of image obtained (surface versus cross section; Marsh et al., 1991).

In this study, a standardized method of determining maximum cranial width and length in individuals with sagittal synostosis is presented. The ability of the traditional CI (using glabella to opisthocranium as the lengthwise measure) to distinguish children with ISS and children with normal skull shape was assessed by using volumetric reconstruction of the head both including and excluding the overlying skin. In addition, the sensitivity and specificity of the traditional CI was compared with alternative scaphocephaly severity indices (SSIs) by using 3D CT surface and internal landmarks referenced to the plane of the skull base.

## MATERIALS AND METHODS

### Patient Sample and Data Acquisition

The sample population consisted of 60 CT head scans from children with ISS (47 boys and 13 girls) and scans from 41 age-matched controls (24 boys and 17 girls scanned for non-head-shape-related indications and determined to have a normal study) who were seen at the Children's Hospital Medical and Regional Center in Seattle, WA. The population age frequencies are shown in Figure 2. The median age was 8 months for patients with ISS and 12 months for patients with normal head shape. Computed tomography data were acquired with a

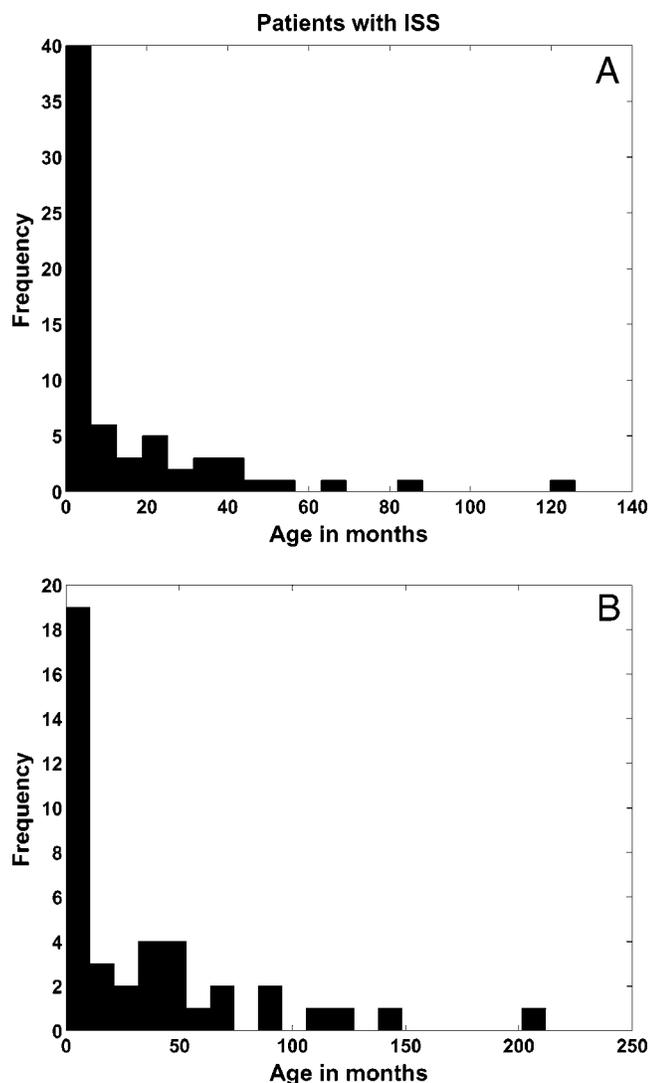


FIGURE 2 Age frequencies for study population of (a) patients affected with ISS and (b) patients with normal head shapes.

high-speed multidetector system (Toshiba Aquilion-16, Toshiba America Inc., Tustin, CA), which produces true isotropic 3D images with a resolution of 0.5 mm. Three-dimensional reformations of each patient's head with soft tissue (to include skin) and bone algorithms were computed with a high-performance viewer-analyzer running on a Vitrea workstation (Toshiba America Inc.) to generate 3D surface meshes (Fig. 1).

### Methods of Measurement

Three-dimensional reformations of the head and skull were used to compute traditional CIs and five new SSIs (SSI- $S_{\text{skin}}$ , SSI- $S_{\text{bone}}$ , SSI-A, SSI-F, and SSI-M) for each patient. Measurements were performed as follows.

*CI.* The ratio of the cranial width to the cranial length is used as the standard definition of CI (Kolar et al., 1997). The cranial length is defined as the distance from the glabella to the opisthocranium (G-OP), and the cranial width is defined

as the distance from euryon to euryon (EU-EU). Figure 3 illustrates an example of CI measurement. The cranial length ( $\alpha$ ) was computed from a calibrated lateral view of the head represented as a 3D surface mesh (Fig. 3a). In a similar manner, a top view of the surface mesh was used to measure the cranial width ( $\beta$ ) (Fig. 3b). The CI was computed as  $CI = \beta:\alpha$ .

*SSIs in the S-Plane (Skull Base Approximated by Using Readily Identifiable External Anatomic Features).* With a calibrated lateral view of CT surface mesh, a 2D plane was traced intersecting the glabella and the meatus of the external auditory canal. This plane was shifted superiorly until it intersected the inion (the most prominent point of the occipital protuberance of the skull). The resulting plane was called the S-plane (Fig. 3c and 3d), which was used to compute the SSI- $S_{\text{bone}}$  and SSI- $S_{\text{skin}}$ . The SSI- $S_{\text{bone}}$  was defined as the ratio of the cranial width to the cranial length as measured on the CT bone slice taken at the level of the S-plane,  $SSI-S_{\text{bone}} = \beta:\alpha$  (Fig. 4a). The SSI- $S_{\text{skin}}$  was also computed as a  $\beta:\alpha$  ratio, but instead of using a CT bone image, soft tissue images were used to include overlying skin to make our measurements (Fig. 4b and 4c).

*SSIs in the A-, F-, and M-Planes (Skull-Base Plane Determined With Distinct Bony Landmarks; Planes for Analysis Determined With Distinct Internal Landmarks).* With a lateral view of a 3D reformation of the skull, a skull-base plane was determined and three parallel planes were subsequently generated to three distinct internal anatomical landmarks based on cerebral ventricles. The skull-base plane was determined by using the frontal nasal suture anteriorly and the opisthion (the middle point on the posterior margin of the foramen magnum of the occipital bone of the skull) posteriorly. This plane was shifted superiorly until positioned (1) just above the top of the lateral ventricles (A-plane), (2) at the Foramina of Munro (F-plane), and (3) at the level of the maximal dimension of the fourth ventricle (M-plane). Computed tomography slices taken at the level of these planes were used to calculate three  $\beta:\alpha$  ratios, which defined the severity indices SSI-A, SSI-F, and SSI-M, respectively (Fig. 5).

### Methods of Analysis

#### Receiver Operating Characteristic Data Analysis

Receiver operating characteristic (ROC) analysis is a graphical method used to evaluate the accuracy of a diagnostic test as a trade-off between its sensitivity and specificity. The sensitivity ( $sn$ ) is the proportion of patients with disease who test positive. The specificity ( $sp$ ) is the proportion of patients without disease who test negative. The key ideas about ROC analysis can be illustrated by means of a simple example that uses CI values. Consider the graph in Figure 6 showing the number of patients with and without ISS arranged according to the value of CI. These distributions overlap; that is, the test does not distinguish normal from disease with 100% accuracy. The area of overlap indicates where the test cannot distinguish normal from synostotic head shape. In practice, for diagnostic

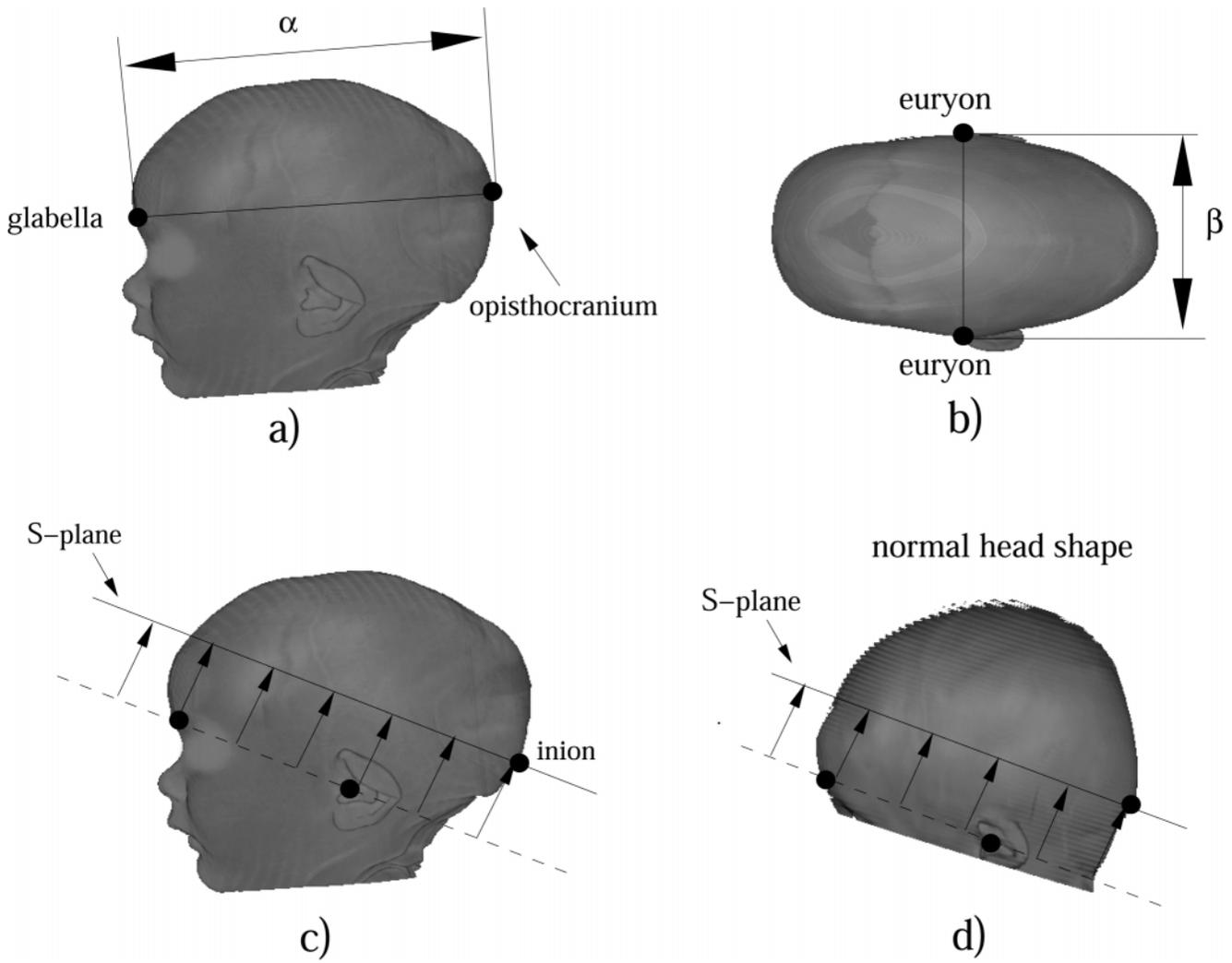


FIGURE 3 The CI was computed by measuring the head (a) length  $\alpha$  and (b) the head width  $\beta$  on a surface mesh representing the skin of a patient's head. Construction of the S-plane on (c) a patient affected with ISS and (d) a patient with normal head shape.

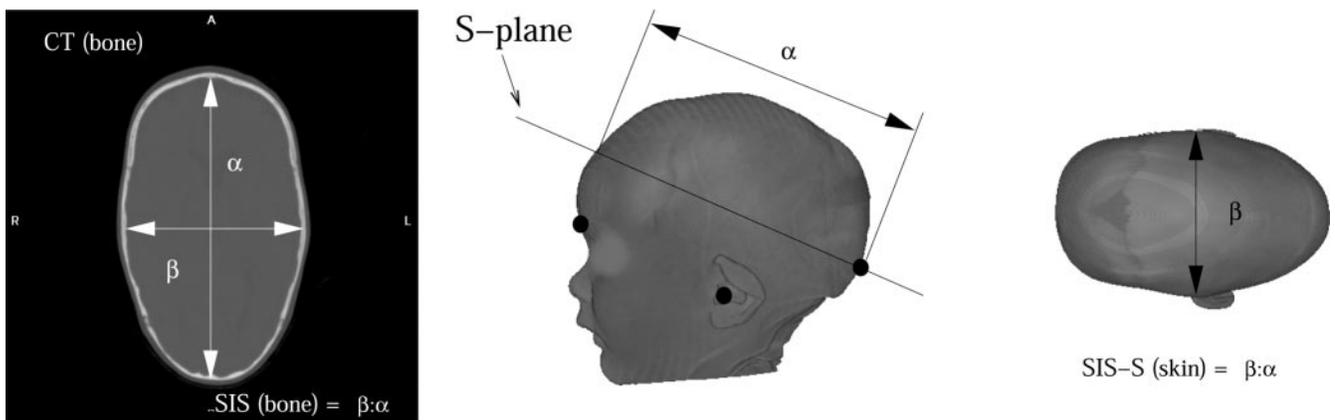


FIGURE 4 (a) Computation of the  $SSI-S_{bone}$  as the ratio  $\beta:\alpha$ . Computation of the  $SSI-S_{skin}$  after identification of (b) the S-plane and head length and (c) head width from skin-surface mesh.

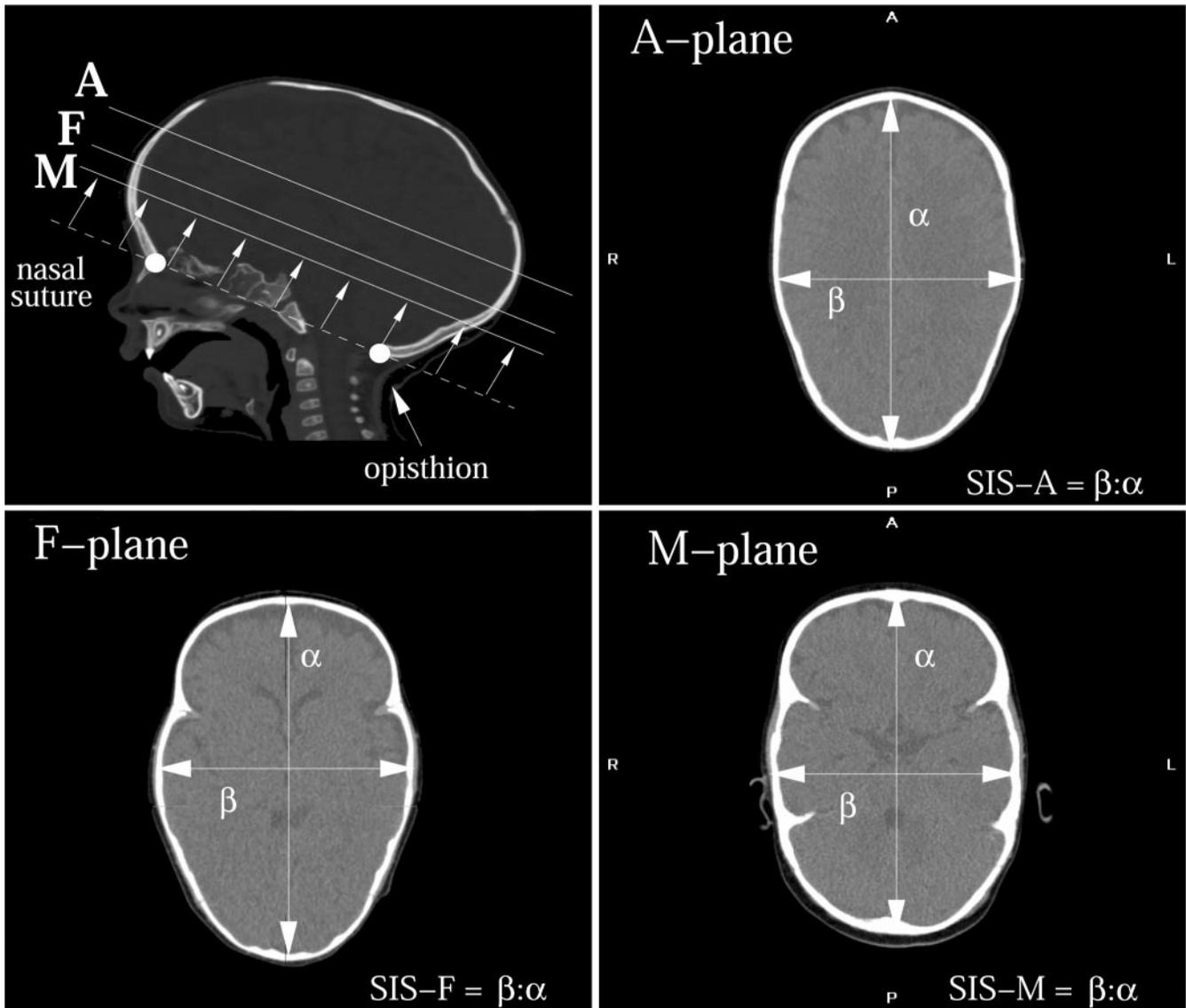


FIGURE 5 SSI-A, SSI-F, and SSI-M were computed from CT bone slices defined by internal anatomical landmarks based on cerebral ventricles just above the top of the lateral ventricles (A-plane), the Foramina of Munro (F-plane), and at the level of the maximal dimension of the fourth ventricle (M-plane).

tests based on quantitative criteria, a threshold is often chosen below which the test is considered to be positive (e.g., abnormal CI) and above which the test is considered to be negative (e.g., normal CI). The choice of the threshold clearly influences the true- and false-positive rates of the diagnostic test. In the context of the current study, the choice of CI threshold influences the proportion of patients to test positive given that they have synostosis (sensitivity) and the proportion of patients

to test negative given that they have normal head shapes (specificity). In practice, the cost of false positives are rarely the same as false negatives, and thus an individual might wish to use different thresholds for different clinical situations to minimize one of the erroneous types of results. A standard way of choosing a threshold is to construct a graph of  $sn$  versus  $sp$  (curve B in Fig. 6). The area under the curve (AUC) is a measure of test accuracy. An area of 1.0 represents a perfect

TABLE 1 Correlation Coefficient Comparing CI Versus SSIs for the Population of Patients With ISS and Patients with Normal Head Shape

Population	SSI-A	SSI-F	SSI-M	SSI-S <sub>bone</sub>	SSI-S <sub>skin</sub>
ISS	0.73	0.83	0.73	0.82	0.81
Normal head shape	0.87	0.75	0.82	0.78	0.76

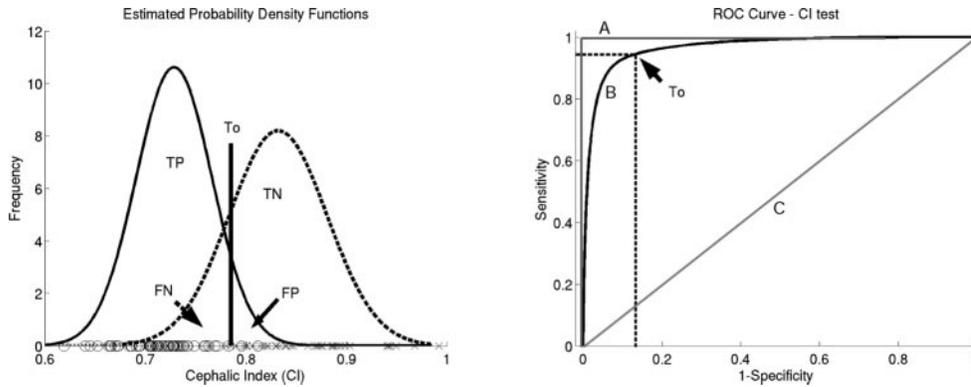


FIGURE 6 (a) Frequency distributions of CI for affected (circles) and unaffected (crosses) patients. (b) ROC curve.

test (curve A in Fig. 6). An area of 0.5 represents a test that is no better than a coin flip (curve C in Fig. 6).

**Density Estimation.** A principled approach to compute a smooth ROC curve from a finite data sample consists of estimating the probability density functions (pdfs) of test values associated with the affected patients and patients with normal head shapes. A parametric method was used for estimating density, in which all the pdfs were modeled as mixtures of normal densities (MONs) (Figueiredo and Jain, 2002). Mixtures of this kind can be used to describe any complex pdf (see Appendix 1). In this work, the parameters of the mixtures were generated by applying the expectation maximization (EM) algorithm (Dempster et al., 1977; McLachlan and Krishnan, 1997), which is commonly used in science and engineering applications. A density estimation example of CI test values for patients with ISS (solid line) and unaffected patients (dashed line) is shown in Figure 6. Samples are shown on the horizontal axis as circles (patients with ISS) and crosses (patients with normal head shapes). The basics of the EM methodology for MONs are beyond the scope of this study; the reader is referred to Appendix 2 for details.

**Confidence Intervals for ROC Curves.** To compare the ROC curves for CI test with the curves of our proposed severity indices (SSI-S<sub>skin</sub>, SSI-S<sub>bone</sub>, SSI-A, SSI-F, and SSI-M), confidence intervals of the curves were estimated by bootstrap. The bootstrap is a nonparametric technique for assigning measures of accuracy to statistical estimates; that is, the resampling method does not rely on assumptions about the underlying distributions among affected and unaffected individuals (Efron, 1982; Efron and Tibshirani, 1993). Appendix 2 describes

an algorithm for estimating confidence intervals of ROC curves. With a similar algorithm, bootstrap confidence intervals were calculated for the accuracy (AUC) of CI and the SSIs, as well as specificity confidence intervals for a given set sensitivity values.

## RESULTS

### Correlation Analysis

A correlation analysis of CI and all SSIs was performed. There was a low-moderate correlation ( $0.75 < p < .85$ ) between CI and each of the SSIs (Table 1). Scatter plots of CI versus SSI-A and CI versus SSI-S<sub>skin</sub> showed a low overlap between the two populations (Fig. 7). Scatter plots for the other SSIs were similar, explaining why all SSIs and CI have high accuracy.

### ROC Curves Accuracy

The shape of the ROC curve of SSI-A in Figure 8 resembles the shape of an ideal test (i.e., one that starts at the origin and steps to a sensitivity of 1 for all nonzero values of the false-positive rate), indicating that this index was more accurate than the others. Note, however, that all five SSIs had better average accuracy than CI (Table 2). In particular, the SIA-A was 3% more accurate on average than CI, and SSI-S<sub>bone</sub> and SSI-S<sub>skin</sub> were 1.5% more accurate than CI. Note that the SSI-S<sub>skin</sub> and SSI-S<sub>bone</sub> had comparable performance.

### ROC Curves Sensitivity and Specificity

The mean specificity for SSI-A was 96% at a sensitivity level of 95%. This means that by using SSI-A to predict ISS, the proportion of patients to test positive given that they have ISS is 95%, and the proportion of patients to test negative given that they have normal head shape is 96% (Table 3 columns C and D). This number contrasts with the 85% specificity of CI. Specificity differences between SSI-A and traditional CI were more dramatic at a sensitivity level of 98% (Table 3 columns E and F). For instance, the SSI-A had a specificity

TABLE 2 Mean Test Accuracy and Their Respective Confidence Intervals for All SSIs and the CI

Index	Mean Accuracy (AUC)*	Confidence Interval
CI	0.9656	(0.9604, 0.9702)
SSI-S <sub>skin</sub>	0.9818	(0.9776, 0.9974)
SSI-S <sub>bone</sub>	0.9839	(0.9860, 0.9811)
SSI-A	0.9917	(0.9892, 0.9932)
SSI-F	0.9866	(0.9838, 0.9919)
SSI-M	0.9816	(0.9783, 0.9871)

\* AUC = area under the curve.

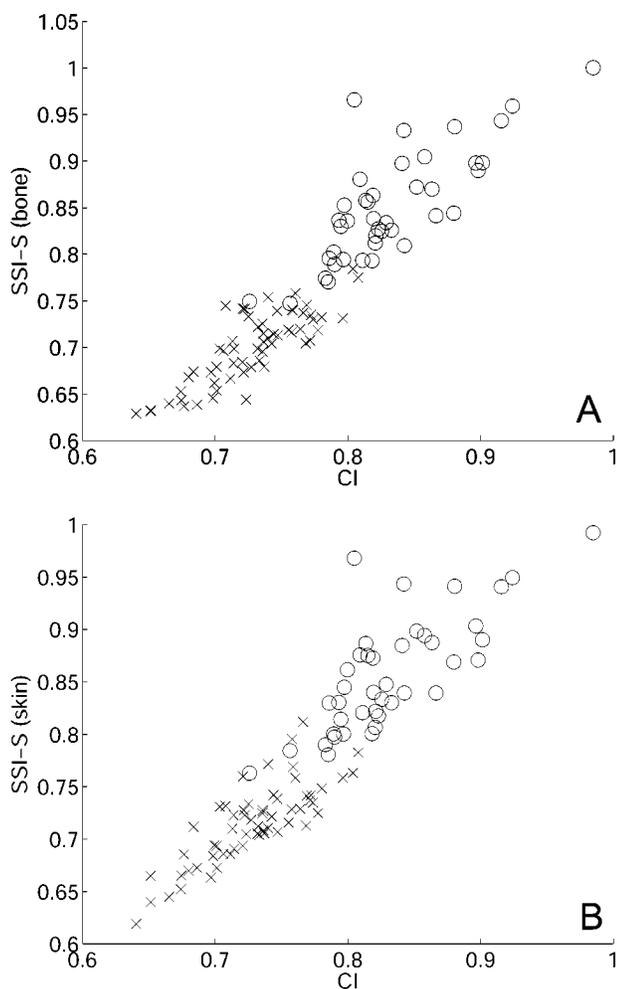


FIGURE 7 Scatter plots (a) CI versus SSI-A and (b) CI versus SSI-S<sub>skin</sub>.

of 94% in contrast with the 68% of CI (a 26% predictive performance difference). In general, it was observed that the specificity of CI was below 87% for sensitivity levels above 93%. Finally, note that the sensitivities of the SSI-S<sub>skin</sub> and SSI-S<sub>bone</sub> were comparable (Fig. 9).

### DISCUSSION AND FUTURE WORK

The traditional CI in children with ISS is generally decreased as compared with children with normal head shape. In practice, however, there is much variation in the methodology in determining cranial width and cranial length, making CI

measurements across different studies difficult to compare. This study evaluated a method to measure CI by using traditional craniometric landmarks of the glabella to opisthocranium for the maximal length of the skull and showed that its average accuracy was 96%. However, it was not specific enough for sensitivity levels above 93%. It was noted that the ratio of cranial width and length measured in a plane parallel to the skull base provided improved sensitivity and specificity for sagittal synostosis. Furthermore, five SSIs computed from planes precisely defined in terms of internal or external cranial anatomy were described. Among these indices, the SSI-A performed the best, which was computed on a plane parallel to the skull base at a level just above the top of the lateral ventricles. This plane captured both the frontal bossing and the inferior projection, or bossing, of the occiput.

The specificity and sensitivity of the SSI-S values was also compared by using external landmarks obtained from reconstructed skin and bone-surface meshes of 3D CT scans. Measurements obtained from skin and bone-surface meshes were comparable. Both were very sensitive and specific. These data pave the way for future comparative studies between CT scan data and surface meshes of cranial form derived from 3D camera systems.

Although the SSI measurements outperformed the CI quantifications, additional quantitative measurements of head shape are needed to capture shape information not detected by the ratio of width to length (Richtsmeier et al., 1998; Lale and Richtsmeier, 2001). It is clear from subjective evaluation of head shapes of patients with sagittal synostosis that a broad range of variations in frontal bossing, occipital bossing, and central narrowing exists. These variations do not appear to be directed by the origin or extent of suture fusion and may be of fundamental interest in understanding the pathogenesis and clinical course of these patients. New techniques are currently being developed to quantify synostotic head shape to address this issue. Three-dimensional photographs generated from a multicamera system are also being used to construct surface meshes and compute SSI-S<sub>skin</sub> surface indices by using external landmarks.

Despite the encouraging results, we acknowledge that a larger sample size is required to achieve more statistically significant estimations. For this reason, we are currently expanding our population of sagittal and control patients. We also note that future comparisons of our proposed indices and the CI will add to the validity of their use as a clinical evaluation tool of sagittal synostosis.

TABLE 3 Mean Specificity at Sensitivity Levels of  $sn = 90\%$ ,  $95\%$ , and  $98\%$  (Columns A, C, and E, respectively) for all SSIs and the CI; Confidence Intervals at  $p = 0.05$  Significance Level (Columns B, D, and F)

Index	A	B	C	D	E	F
CI	0.9279	(0.9140, 0.9464)	0.8499	(0.8271, 0.8761)	0.6851	(0.6527, 0.7156)
SSI-S <sub>skin</sub>	0.9678	(0.9479, 0.9782)	0.9430	(0.9091, 0.9754)	0.8917	(0.8679, 0.9079)
SSI-S <sub>bone</sub>	0.9602	(0.9528, 0.9656)	0.9398	(0.9300, 0.9474)	0.9066	(0.8928, 0.9173)
SSI-A	0.9799	(0.9755, 0.9837)	0.9676	(0.9578, 0.9736)	0.9462	(0.9264, 0.9573)
SSI-F	0.9600	(0.9523, 0.9715)	0.9396	(0.9283, 0.9515)	0.9066	(0.8896, 0.9218)
SSI-M	0.9504	(0.9371, 0.9740)	0.9058	(0.8838, 0.9385)	0.8472	(0.8265, 0.8645)

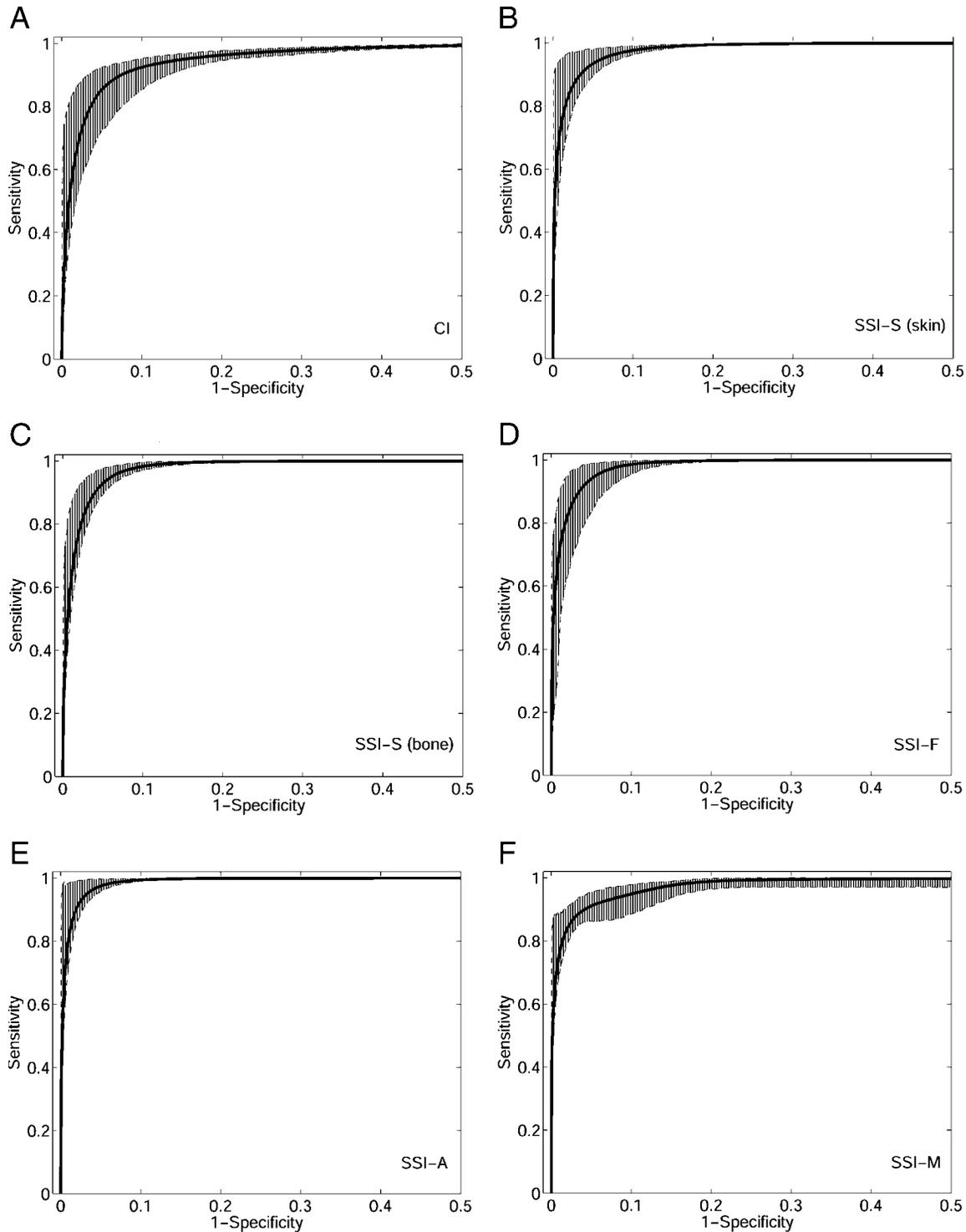


FIGURE 8 ROC curves and  $p = .05$  confidence intervals for CI, SSI-S<sub>skin</sub>, SSI-S<sub>bone</sub>, SSI-F, SSI-A, and SSI-M.

We recognize the existence of alternative density estimation methods that could be used to generate smooth ROC curves (e.g., Silverman, 1986; Zou et al., 1998; Sorribas et al., 2002; Zhou and Harezlak, 2002), particularly nonparametric approaches such as kernel density estimation (Silverman, 1986; Zou et al., 1998). A comparative study of parametric and non-

parametric methods for estimating ROC curves in the context of this work as well as comparisons with alternative severity measures such as intracranial volume are topics of a future investigation.

Finally, we acknowledge that our proposed methodology requires state-of-the-art radiological resources that are not al-

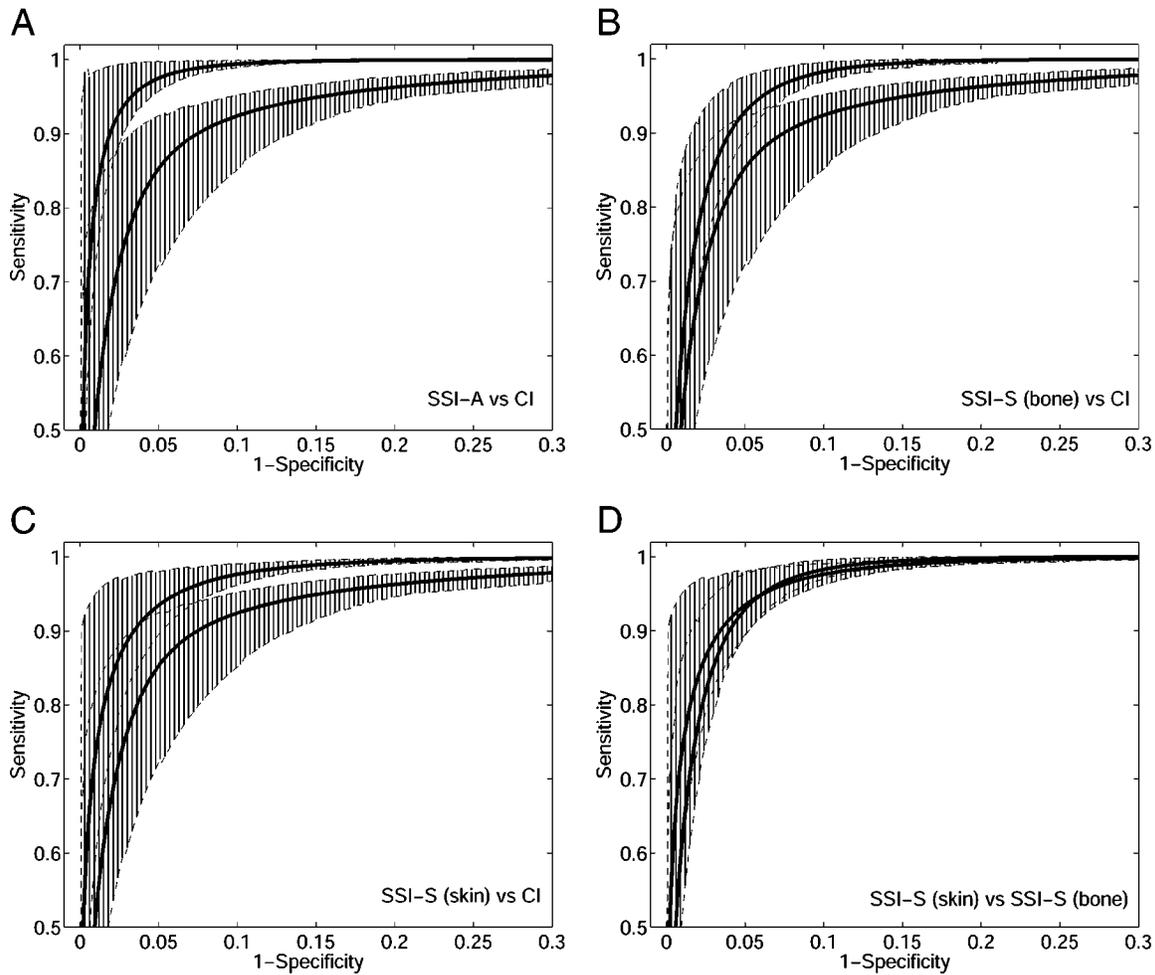


FIGURE 9 ROC curves and  $p = .05$  confidence intervals for SSI-A versus CI, SSI-S<sub>bone</sub> versus CI, SSI-S<sub>skin</sub> versus CI, and SSI-S<sub>skin</sub> versus SSI-S<sub>bone</sub>.

ways available to clinicians. However, as pointed out in the introduction, there has been considerable well-documented work on quantitatively analyzing head shape and surgical outcomes in sagittal synostosis by simple, 2D CT scan techniques that can be alternatively considered in clinical settings.

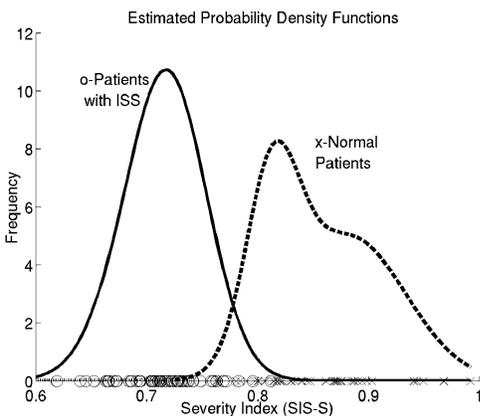


FIGURE 10 Mixture of normal models for SSI-S<sub>skin</sub> for normal (dashed line) patients and patients with ISS (solid line).

## REFERENCES

Albright AL, Towbin RB, Schultz BL. Long term outcome after sagittal synostosis operations. *Pediatr Neurosurg.* 1996;25:78–82.

Christophis P, Junger TH, Howaldt HP. Surgical correction of scaphocephaly: experiences with a new procedure and follow up investigations. *J Maxillofac Surg.* 2001;29:33–38.

Cohen MM, MacLean RE. *Craniosynostosis. Diagnosis, Evaluation and Management.* New York: Oxford University Press; 2000.

David DJ, Simpson DA. Surgical indications in single suture craniosynostosis. *Neurosurgery.* 1982;11:466.

David L, Proffer P, Hurst WJ, Glazier S, Argenta LC. Spring-mediated cranial reshaping for craniosynostosis. *J Craniofac Surg.* 2004;15:810–816.

Dempster AP, Laird AM, Rubin DB. Maximum likelihood from incomplete data via the em algorithm. *J R Stat Soc.* 1977;39:1–38.

Efron B. *The Jackknife, the Bootstrap, and Other Resampling Plans.* Philadelphia: Society for Industrial and Applied Mathematics; 1982.

Efron B, Tibshirani RJ. *An Introduction to the Bootstrap.* London: Chapman & Hall; 1993.

Farkas L, Hreczko TA, Katic MJ. Craniofacial norm in North America, caucasians from birth (one year) to young adulthood. In: Farkas LG, ed. *Anthropometry of the Head and Face in Medicine.* 2nd ed. New York: Raven Press; 1994:241–336.

Fata JJ, Turner MS. The reversal exchange technique of total calvarial reconstruction for sagittal synostosis. *Plast Reconstr Surg.* 2001;107:1637–1646.

Figueiredo M, Jain AK. Unsupervised learning of mixture models. *IEEE PAMI*. 2002;3:381–396.

Friede H, Lauritzen C, Figueroa AA. Roentgencephalometric follow up after early osteotomies in patients with scaphocephaly. *J Craniofac Surg*. 1996;7:96–101.

Guimaraes-Ferreira J, Gewalli F, David L, Olsson R, Friede H, Lauritzen CGK. Clinical outcome of the modified pi-plasty procedure for sagittal synostosis. *J Craniofac Surg*. 2001;12:218–224.

Guimaraes-Ferreira J, Gewalli F, David L, Olsson R, Friede H, Lauritzen CGK. Spring-mediated cranioplasty compared with the modified pi-plasty for sagittal synostosis. *Scand J Plast Reconstr Surg Hand Surg*. 2003;37:208–215.

Hunter AG, Rudd NL. Craniosynostosis. I. Sagittal synostosis: its genetics and associated clinical findings in 214 patients who lacked involvement of the coronal suture(s). *Teratology*. 1976;14:185–193.

Jimenez DF, Barone CM, McGee ME, Cartwright CC, Baker CL. Endoscopy-assisted wide vertex craniectomy, barrel stave osteotomies, and postoperative molding therapy in management of sagittal suture craniosynostosis. *J Neurosurg (5 Suppl Pediatrics)*. 2004;100:407–417.

Kaiser G. Sagittal synostosis—its clinical significance and the results of three different methods of craniectomy. *Childs Nerv Syst*. 1988;4:223–230.

Kolar JC, Salter EM, Charles C. *Craniofacial Anthropometry, Practical Measurement of the Head and Face for Clinical, Surgical and Research Use*. Springfield, IL: Charles C. Thomas; 1997.

Lajeunie E, Le Merrer M, Bonaiti-Pellie C, Marchac D, Renier D. Genetic study of scaphocephaly. *Am J Med Genet*. 1996;62:282–285.

Lale SR, Richtsmeier JT. *An Invariant Approach to Statistical Analysis of Shapes*. Boca Raton, FL: Chapman & Hall; 2001.

Marsh MD, Jenny A, Galic M, Selwyn P, Vannier MW. Surgical management of sagittal synostosis. *Neurosurg Clin N Am*. 1991;2:629–640.

McLachlan GJ, Krishnan T. *The EM Algorithm and Extensions*. Hoboken, NJ: John Wiley and Sons; 1997.

Panchal J, Marsh JL, Park TS, Haufman B, Pilgram T, Huang SH. Sagittal craniosynostosis outcome assessment for two methods and timings of intervention. *Plast Reconstr Surg*. 1999;103:1574–1999.

Posnick JC, Armstrong D, Bite U. Metopic and sagittal synostosis: intracranial volume measurements prior to and after cranio-orbital reshaping in childhood. *Plast Reconstr Surg*. 1995;96:299–309.

Posnick JC, Bite U, Nakano P, Davis J, Armstrong D. Indirect intracranial volume measurements using CT scans: clinical applications for craniosynostosis. *Plast Reconstr Surg*. 1992;89:34–45.

Posnick JC, Lin KY, Chen P, Armstrong D. Sagittal synostosis: quantitative assessment of presenting deformity and surgical results based on CT scans. *Plast Reconstr Surg*. 1993;92:1015–1024.

Richtsmeier JT, Cole TM, Krovitz G, Valerie CJ, Lele S. Preoperative morphology and development in sagittal synostosis. *J Craniofac Genet Dev Biol*. 1988;18:64–78.

Seymour-Dempsey K, Baumgartner JE, Teichgraeber JF, Xia JJ, Waller AL, Gateno J. Molding helmet therapy in the management of sagittal synostosis. *J Craniofac Surg*. 2002;13:631–635.

Silverman BW. *Density Estimation for Statistics and Data Analysis*. New York: Chapman & Hall; 1986.

Slomic AM, Bernier JP, Morissette J, Reinier D. A craniometric study of sagittal craniosynostosis (SC). *J Craniofac Genet Dev Biol*. 1992;12:49–54.

Solan GM, Wells KC, Raffel C, McComb JG. Surgical treatment of craniosynostosis: outcome analysis of 250 consecutive patients. *Pediatrics*. 1997;100:E2.

Sorribas A, March J, Trujillano J. A new parametric method based on S-distributions for computing receiver operating characteristic curves for continuous diagnostic tests. *Stat Med*. 2002;21:1213–1235.

Waitzman AA, Posnick JC, Armstrong DC, Gaylene EP. Craniofacial skeletal measurements based on computed tomography: part II. Normal values and growth trends. *Cleft Palate Craniofac J*. 1992;29:118–128.

Zhou XH, Harezlak J. Comparison of bandwidth selection methods for kernel smoothing of ROC curves. *Stat Med*. 2002;21:2945–2955.

Zou KH, Tempany CM, Fielding JR, Silverman SG. Original smooth receiver operating characteristic curve estimation from continuous data: statistical

methods for analyzing the predictive value of spiral CT of uretral stones. *Acad Radiol*. 1998;5:680–687.

## APPENDIX 1

### Mixtures of Normal Densities

MONs are combinations of normal distributions or, more specifically, a weighted sum of individual normal densities. Recall that the one-dimensional pdf of a normal density with mean  $\mu$  and variance  $\sigma$  is defined as

$$f(x; \mu, \sigma) = 1/(2\pi\sigma^2)^{0.5} \exp\{-(x-\mu)^2/\sigma^2\}.$$

A weighted combination of  $K$  densities of the form

$$g(x) = \lambda_1 f(x; \mu_1, \sigma_1) + \dots + \lambda_k f(x; \mu_k, \sigma_k)$$

such that  $\lambda_1 + \dots + \lambda_k = 1$  is defined as the pdf of a MON. By varying the number of components  $K$ , the weights  $\lambda_i$ , and the parameters  $\mu_i$  and  $\sigma_i$  of each function, it is possible to describe any complex pdf (Figueiredo and Jain, 2002). For instance, the pdf of the SSI- $S_{\text{skin}}$  for patients with normal head shape (dashed line) shown in Figure 10 is a mixture

$$g(x) = 0.3732f(x; 0.8133, 5.7 \times 10^{-4}) \\ + 0.626f(x; 0.8831, 0.3732).$$

In practice, the EM algorithm is used to learn the parameters of Gaussian mixtures from a finite data sample. The reader is invited to consult Dempster et al. (1977), McLachlan and Krishnan (1997), and Figueiredo and Jain (2002) for a comprehensive description of the EM and model selection techniques that were used in this work.

## APPENDIX 2

### Bootstrap Confidence Intervals

A bootstrap-based algorithm for computing a confidence interval for a ROC curve is summarized as follows (Efron and Tibshirani, 1993). The input is a set of severity index values computed from a population of size  $N$  of affected and unaffected patients. The output is a mean ROC curve and its corresponding confidence interval.

1. Obtain a random sample of test values with replacement of size  $n < N$  from the population of affected and unaffected individuals, and estimate a ROC curve by using the methods described in the “Materials and Methods” section.
2. Repeat step 1 for  $L$  times, where  $L$  is sufficiently large (i.e., from 200 to 10,000).
3. For each false-positive rate  $fp = 1 - sp$  ( $sp = \text{specificity}$ ) in the interval  $[0,1]$ . Create an array  $[sn]$  of size  $L$  that stores the sensitivity values of the curves computed in step 2 that correspond to the given specificity value  $sp$ . Compute the mean sensitivity value  $SN$  for the values in  $[sn]$ .
  - Create an array  $[d]$  that stores the differences between the

elements of  $[sn]$  and the mean SN; that is,  $d_i = SN - sn_i$  for  $i = 1, \dots, L$ .

- Create two arrays  $[a]$  and  $[b]$ . The array  $[a]$  contains all the differences of array  $[d]$  that are negative. The array  $[b]$  contains all the differences of array  $[d]$  that are positive.
- Compute the  $p$ -confidence interval  $[SN, SN + b]$  for the

positive differences in  $[b]$ ; that is, find  $b$  such that  $P(|SN - sn| < b) = p$ , where  $P(\cdot)$  denotes the probability of an event and  $p$  is the significance level.

- Compute the  $p$ -confidence interval for the negative differences  $[SN - a, SN]$ .
- Compute the confidence interval of SN as  $[SN - a, SN + b]$ .