Predicting Neuropsychological Development from Skull Imaging

H. J. Lin, S. Ruiz-Correa, L. G. Shapiro, M. L. Speltz, M.L. Cunningham and R. W. Sze

Abstract— **Craniosynostosis is a serious and common pediatric disease caused by the premature fusion of sutures of the skull. Although studies have shown an increase in risk for cognitive deficits in patients with isolated craniosynostosis, the causal basis for this association is still unclear. It is hypothesized that an abnormally shaped skull produces a secondary deformation of the brain that results in the disruption of normal neuropsychological development. In this paper, we conduct a comparative analysis of our newly developed shape descriptors in an attempt to understand the impact of skull deformations on neurobehavior. In particular, we show that our *scaphocephaly severity indices* and *symbolic shape signatures* are predictive of mental ability and psychomotor functions, respectively, which suggests the possibility that secondary deformation could influence neuro-developmental status.**

I. Introduction

Craniosynostosis is the pathological condition of premature fusion of calvarial (skull) sutures or fibrous skull joints in childhood, affecting 1 in 2,500 individuals. Normally, an infant is born with open sutures, allowing for the development and expansion of the brain (Fig. 1a). However, in children with craniosynostosis, one or more of these sutures close prematurely. The early closure of these sutures results in abnormalities in calvarial shape due to the combination of restriction of osseous growth perpendicular to the fused suture and compensatory growth in unfused calvarial bone plates. Isolated sagittal synostosis (ISS), denoted by a long narrow skull shape (Fig. 1b) is the most common form of isolated suture synostosis with an incidence of approximately 1 in 5,000, accounting for 40-60% of single suture synostoses [10]. The degree of skull shape deformity and changes in frontal and occipital bossing and biparietal narrowing can vary significantly between individuals with ISS (Fig. 2).

Although once believed to have negligible neurodevelopmental consequences, single-suture synostosis has been associated in many recent studies with a variety of neuropsychological deficits in infants and children. Speltz et al. suggested that the risk of poor neurobehavioral outcome in children with single-suture synostosis may be three-to-five times higher than average [24]. Magge and colleagues also identified statistical significance deficits in global IQ in children with sagittal synostosis [14]. Virtanen et al. reported that the children with scaphocephaly had mild deficits in

H.J. Lin is a student at the Department of Medical Education and Biomedical Informatics, L. G. Shapiro is with the Faculty of Computer Science and Engineering, and M. L. Speltz is with the Faculty of Psychiatry and Behavioral Sciences at the University of Washington. M. L. Cunningham is affiliated to the Children's Craniofacial Center at Children's Hospital and Regional Medical Center, Seattle WA. S. Ruiz-Correa and R. W. Sze are with the Dept. of Diagnostic Imaging and Radiology at the Children's National Medical Center, Washington, D.C.
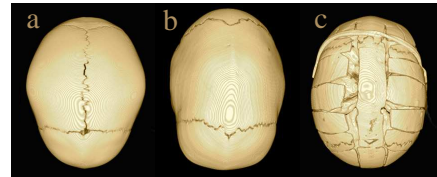

Fig. 1. Three-dimensional reformations from computed tomography imaging showing the top-view of a patient with a) a normal skull shape; b) a malformed skull affected with sagittal synostosis; the sagittal suture is completely absent; and c) a corrected skull after a reconstructive surgery that is often used to palliate the calvarial malformation.

auditory short-term memory and language [25]. Furthermore, Pershing and collegues identified 58% of single-suture craniosynostosis patients as having a learning disability [16]. In summary, it now appears certain that isolated craniosynostosis is associated with elevated risk for mild to moderate neurocognitive impairment [9].

Causal mechanisms that potentially explain this association are poorly understood. Clinical formulations of the potential linkage between synostosis and neurobehavioral functioning have usually assumed a direct, linear pathway in which suture fusion leads to brain deformation and, consequently, neuropsychological deficits [2][19]. However, there are insufficient data and limited theory with which to hypothesize any particular causal pathway. One possible causal factor is secondary cerebral deformation resulting from brain growth in an abnormally shaped skull [1][4]. The possibility of this hypothesis suggests that different skull abnormalities may be associated with cerebral deformation, leading to different levels of severity in neurodevelopmental outcomes.

Fig. 2. Three-dimensional reformations (top-view) showing a variety of calvarial shapes often observed in patients affected with ISS.

In recent years, several attempts have been made to quantify skull morphology in order to perform classification of synostotic shapes. In [11], Lale and Richtsmier combined morphometric techniques, likelihood-based and dissimilarity-based classification methods but reported high classification error rates [11]. Our own scaphocephly severity indices (SSIs), Fourier-based *cranial spectrum* [21] and outline-based *cranial images* [20] have also been developed to quantify skull shapes. In addition, we described in [12] [13] a symbolic signature-based methodology that can accurately and efficiently quantify and predict the most common types of isolated craniosynostosis. Our symbolic shape descriptors (SSDs) represent skull morphology as a blend of *geometric aspects* that are estimated from a *training data* set.

In a first effort to clarify the secondary cerebral deformation hypothesis, we compute nonlinear regression models that attempt to predict neuro-developmental condition (quantified in terms of standard *Bayley scores*), as a function of shape descriptors that have been previously used in the context of a classification task. We show that our SSI and SSD descriptors are predictive of mental and psychomotor performances, respectively. This result suggests that skull morphology in craniosynostosis could be a deleterious factor in the neuropsychological development of affected children.

The task we approach in this paper can be formally described as follows. We are given a random sample of N_M skull shapes as training set. Each shape S in the sample is labeled by a numeric score b that quantifies neurobehavioral status. Using the skull shape information and scores, we wish to construct a regression model that uses shape descriptors to predict the score b_{new} of a new skull shape, S_{new}.

The paper is organized as follows. In sections II.A and B, we briefly describe our methods for acquiring neurobehavioral and imaging data. In section II.C we develop the shape descriptors utilized in our experiments. The nonlinear regression technique used to predict neurobehavior is described in sections II.D and E, and the results of our comparative study are summarized in section III. Section IV concludes our paper.

II. METHODS

A. Quantification of neuropsychological development

Bayley Scales of Infant Development, Edition II (BSIC-II), is used to quantify neuro-developoment on infants with isolated sagittal synostosis who are 30 months of age or younger. This test is designed to assess whether infants are developing at an expected rate in order to plan for intervention strategies [3]. The BSID is a standardized, norm-referenced objective test with mean $\mu = 100$ and standard deviation $\sigma = 15$. It yields separate indices of mental and psychomotor development: the Mental Development Index (MDI) and Psychomotor Development Index(PDI). The MDI score is used to evaluate sensory/perceptual acuities, memory, learning and problem solving, while the PDI assesses degree of body control and coordination of large muscles. The BSID-II continues to be the best and most frequently used instruments to assess infant development in psychology.
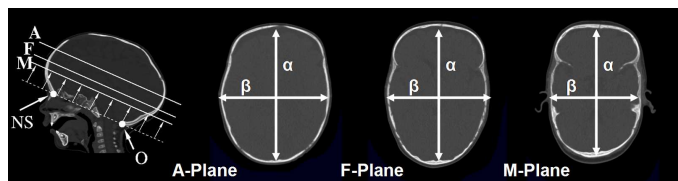

Fig. 3. The scaphocephaly severity indices SSI-A, SSI-F and SSI-M are computed as the head width to head length ratio β/α as measured on CT bone slices that are defined by internal anatomical landmarks on cerebral ventricles.

B. Source of images

Shape indices of the same ISS patients are extracted from computed tomography (CT) image slices from skull imaging. In order to standardize our measurements, we use a calibrated lateral view of a 3-D reconstruction of the skull to select three CT slice planes defined by internal brain landmarks. These planes are parallel to the skull base plane, which is defined by using the frontal nasal suture anteriorly and opsithion posteriorly. The A, F, and M planes are shown in Figure 3. The A-plane is at the top of the lateral ventricle, the F-plane is at the Foramina of Munro, and the M-plane is at the level of the maximal dimension of the fourth ventricle. Using standard image segmentation and spline interpolation techniques [7], it is possible to extract the oriented outline from a CT bone image at the level of any of the planes defined above.

C. Shape descriptors

1) *Scaphocephaly severity indices* SSI-A, SSI-F and SSI-M are computed as the head width to head length ratio (β/α) measured on the planes A, F and M, respectively (Fig. 3). SSIs were developed in [22] to improve traditional *cephalic index* quantifications. The cephalic index also takes the ratio of head width to head length, but is measured by hand over the skin of the skull. The SSIs assume that a skull shape can be approximated as an ellipse that has eccentricity

$$e = \sqrt{1 - \frac{\beta^2}{\alpha^2}}.$$

Note that if $e = 0$ (that is $\beta/\alpha = 1$) the outline shape would be a perfect circle. An eccentricity value close to one would suggest a very narrow outline shape.
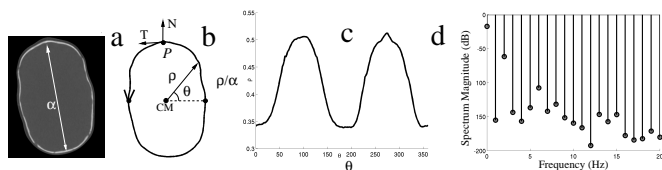

Fig. 4. a) Bone CT slice at the level of the A-plane; b) oriented outline counter clockwise direction; c) same outline represented in polar coordinates (ρ, θ); and d) 21 components of the corresponding cranial spectrum. Key: α (maximum outline length), T (tangent vector), N (normal vector), and (CM) center of mass.

2) *Polar representation (PR)* describes skull shape using a polar coordinate syststem as follows. An oriented outline

is first discretized into N_V evenly spaced vertices. With the origin at the center of the mass of the outline, the angle and the axial distance from the origin to each of the defined vertices on the outline is computed (Fig. 4c). This representation encompasses shape information of the entire contour that cannot be captured by the SSI ratios. In this paper, we compute the polar representation from outlines on the A, F, and M planes previously defined.

Fig. 6. Symbolic labels are assigned to the vertices of oriented outlines by applying k-means clustering to their numeric attributes. Local geometric properties are represented by making strings of these symbols. The set of words is the BOW representation of the skull shape. Oriented outlines of a) sagittal and b) normal skull shape contours, computed at the level of the A-plane. The size of the strings is $N_W = 3$.
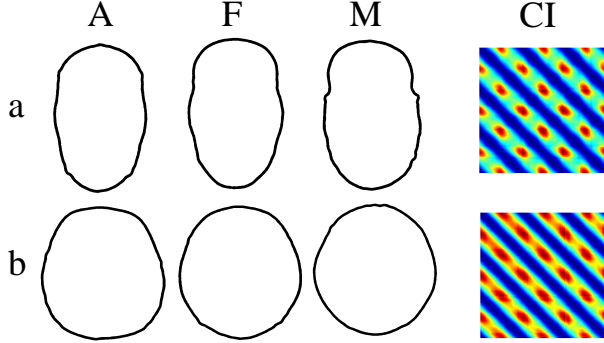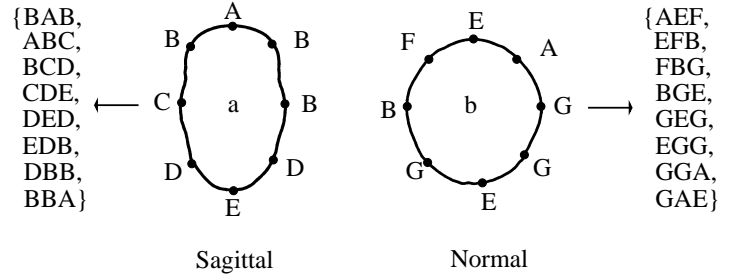

Fig. 5. Cranial image (CI) descriptors for the oriented outlines computed on the A, F and M planes. a) Outlines of a skull affected by ISS. b) Outlines of a normal skull shape.

3) *Cranial spectrum* (CS) is a Fourier-based shape representation developed in [20] that describes a skull shape with the magnitude of the Fourier series coefficient of a periodic function. This function is derived from a normalized oriented outline by using polar coordinates described in the previous section (Fig. 4). This shape representation is closely related to traditional DFT-based descriptors [18]. We use the first N_R coefficients of the spectrum from each of the three planes together in our experiments.

4) *Symbolic shape descriptors* (SSDs) were developed in [12][13] to reduce the dimensionality of numeric shape descriptors, such as the *cranial image* [20], while preserving their descriptive power. The cranial image (CI) is a dense matrix representation of pairwise normalized square distances. The matrices are computed for all the vertices of N_L oriented outlines (such as the ones computed on the A, F and M planes) that have been discretized to N_V evenly spaced vertices (Fig. 5).

For the sake of completeness, the algorithm for computing SSD is briefly summarized here. The interested reader is invited to consult [12] and [13] for details.

4.1) *SSD Computation*. The input are training labeled skull shapes $\mathcal{S}=\{S_1, \cdots, S_{N_M}\}$, Each shape is represented by N_L oriented outlines, and each outline is discretized into N_V vertices. For the sake of simplicity and without loss of generality, we assume that $N_L = 1$. The SSDs are computed as follows:

- For each shape S_j in $\mathcal{S}$ and each vertex v_i of S_j, compute the vector of distances from all other vertices of S_j to v_i. This vector is the same as the i-th row of the CI matrix descriptor (Fig. 6).
- Cluster these vectors by k-means clustering with user-selected k and assign each cluster a symbolic label. Each vertex receives the label of its cluster.

- Compute a *bag of words* (*BOW*) representation of the skull outlines in $\mathcal{S}$. More specifically, the symbols associated with the vertices of an oriented outline are used to construct strings of symbols or words. The word size is fixed at some integer $1 \leq N_W \leq N$. For instance, when $N_W = 3$, each word contains three symbols. A *BOW* representation for the outline in Figure 6 (sagittal skull shape) is the unordered set $s_i = \{BAB, ABC, BCD, CDE, DED, EDB, DBB, BBA\}$.
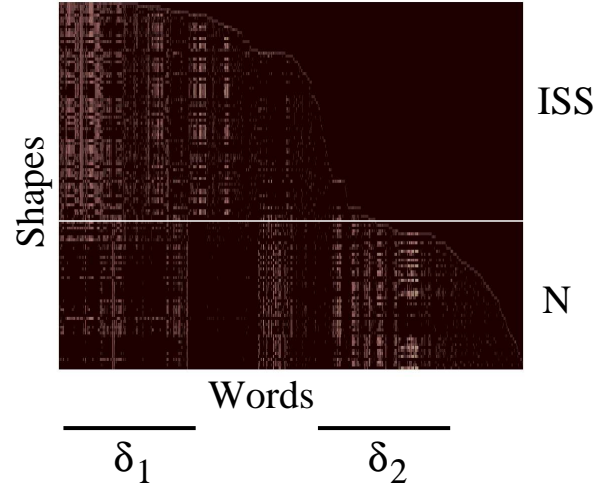

Fig. 7. Co-occurrence distribution of the local geometric aspects corresponding to the skull shapes in our sample population. Each skull shape is represented using the all three A, F, and M planes. Key: Sagittal (S) and Normal (N).

- Compute a $N_M \times N_Q$ co-occurrence matrix of counts $[n(s_i, w_j)]_{ij}$ for the training set where $n(s_i, w_j)$ denotes the number of times the word w_j occurs in the *BOW* s_i associated with the skull outline S_i (Fig. 7). N_Q is the total number of distinctive words computed from the training set.
- Apply *probabilistic latent semantic analysis* (PLSA) to the co-occurrence matrix of the training set [8]. PLSA is a latent variable model which associates an unobserved class variable $z_k \in z_1, \ldots, z_{N_P}$ with each observation, and an observation being the occurrence of a word in a

particular *BOW*. The PLSA is defined as

$$P(s_i, w_j) = P(s_i)P(s_i, w_j) \quad (1)$$

$$P(w_j|s_i) = \sum_{k=1}^{N_P} P(w_j|z_k)P(z_k|s_i), \quad (2)$$

where $P(s_i, w_j)$ denotes the probability that a word occurrence will be observed in a particular *BOW* s_i, $P(w_j|z_k)$ denotes the class-conditional probability of observing the word w_j given the hidden variable z_k, and $P(z_k|s_i)$ denotes a *BOW*-specific probability distribution over the latent variable space.

- Use the class-conditional probabilities $P(s|z)$ estimated in the previous step to construct the symbolic shape descriptors for the outlines in $\mathcal{S}$. More specifically, for each outline S_i in the training set, form its corresponding *symbolic shape descriptor* as the N_P-dimensional vector

$$[P(s_i|z_1), \cdots, P(s_i|z_{N_P})]. \quad (3)$$

The outputs of the training algorithm are the k-means cluster centers, the set of words in the training set (vocabulary), and the parameters of the PLSA model.

An intuitive interpretation for the aspect model can be obtained by observing that the conditional distributions $P(w_j|s_i)$ are convex combinations of the P class conditionals or *aspects* $P(w_j|z_k)$. In our context, words encode local geometric properties of the outline shapes; therefore, the global geometric properties of the outlines in $\mathcal{S}$ are modeled as convex combinations of local *geometric aspects*. Note that $P(s|z) = P(z|s)P(s)/P(z) \propto P(z|s)$, where $P(s)$ is considered as uniformly distributed. Therefore, given z, shapes can be ranked according to $P(z|s)$ [13]. Also note that the CI descriptors require $O(N_V^2)$ bytes of computer storage [12], while the SSDs only require $O(N_P)$. Experimental evidence has suggested that N_P is usually smaller that N_V by an order of magnitude [13].

D. Regression analysis

The general purpose of a regression analysis is to estimate the function dependence of the dependent variable b (the MDI and PDI scores in our case), on a set of independent variables s (the shape descriptors). Early analysis of our data (not reported in this paper) suggested that the relationship between our shape descriptors and neurobehavioral scores was better modeled using nonlinear regression techniques. We utilized off-the-shelf support vector machine (ν-SVM) regression with a Gaussian kernel to create our models [23]. We selected the width of the Gaussian kernel by minimizing the *leave-one-out* (LOO) estimate of the averaged mean square error (AMSE). The algorithm to compute our regression models is as follows.

Training. The input are training labeled skull shapes $\mathcal{S}=\{S_1, \cdots, S_{N_M}\}$ and neurobehavioral scores $\mathcal{B}=\{b_1, \cdots, b_{N_M}\}$ Each shape is represented by N_L oriented outlines, and each outline is discretized into N_V vertices.

- Compute the numeric/symbolic descriptors for the shapes in $\mathcal{S}$. In the case of symbolic descriptors, use the algorithm described in II.C.4.1.
- Select the model of a ν-SVM regression function by minimizing the LOO estimate of the AMSE. This function relates shape descriptors and their corresponding scores in the training set.

The output is a ν-SVM regression function. The testing algorithm for numeric descriptors is straightforward. The algorithm for the symbolic descriptors is as follows.

Testing. The inputs are a new skull shape S_{new}, the parameters of the PLSA model, and a ν-SVM regression function.

1) Use the class-conditional probabilities $P(w|z)$ estimated from the training set to compute $P(s_{new}|z)$ for the test skull S_{new}, and form the symbolic shape descriptor $[P(s_{new}|z_1), \cdots, P(s_{new}|z_{N_P})]$ [13].
2) Use the regression function and the symbolic shape descriptor computed in the previous step to predict the score of S_{new}.

The output is the predicted score b_{new}.

E. Significance test for SVM regression

Normally, standard convergence bounds would be utilized to suggest how well the estimated error from the regression analysis can be generalized. However, because of the small size of our test set in the cross-validation technique, it is typically difficult to obtain low variances for useful bounds to determine how accurate the estimations are to the true expected error [6]. In addition, it requires extensive modeling of the dependence of errors in cross-validation trials to compute a meaningful variance estimation [5]. Therefore, to obtain an accurate estimate on how close the predicted error is to the true expected error of the trained regressor, it is not common practice to use cross-validation methods.

In this paper, we perform permutation tests to assess statistical significance of the regressor and its performance. Even though permutation tests do not provide the statistical bounds of the predicted error, it provides a weaker guarantee, namely the statistical significance, on how likely we were to obtain the observed association and predicted error by chance [5].

Permutation tests were originally developed in statistics for hypothesis testing. In general, the objective of the hypothesis testing is to reject the null hypothesis at a certain level of significance α, which sets the maximal acceptable probability of declaring that the observed associations are real when the null hypothesis is true. The null hypothesis in our study assumes that the relationship between the data, which are the shape descriptors, and the label, which is the MDI or the PDI score, cannot be learned reliably by the regressor used in the training step. The alternative hypothesis is that we can train the regressor with small expected error. For any value of statistic measuring the predicted error, the corresponding *p-value* is the highest level of significance at which the null hypothesis can still be rejected [5].

Let α be an acceptable level of significance, and Π_l be a set of all permutations of indices $1, \ldots, l$, where l is the number of independent examples in the training dataset. The permutation test procedure for a regressor that consists of N_l iterations is defined as follows:

- For i = 1, $\cdots$, N_l
  1) Sample a permutation π_i^m from a uniform distribution over Π_l.
  2) Compute the AMSE e_i^m for this permutation of scores.
- Construct an empirical cumulative distribution $\hat{P}(E \leq e) = \frac{1}{N_l} \sum_{i=1}^{N_l} \Theta(e - e_i^m)$ where Θ is a step-function ($\Theta(x) = 1$, if $x \geq 0$; 0 otherwise).
- Compute the training error for the actual scores, e_0 and its corresponding p-value p_0 under the empirical distribution $\hat{P}$. If $p_0 \leq \alpha$, reject the null hypothesis.

F. Data sample

Our sample population consists of 71 patients affected with ISS and 53 children with normal skull morphology. The control group came from a study that is reported in Ruiz et al. [22]. Neuro-bahavioral data is not available for these individuals. Computed tomography data are acquired with a multi-detector system that produces isotropic 3-D images with 0.5mm resolution. Three-dimensional reconstructions of each patient's skull are generated with a high performance workstation (Fig. 2).
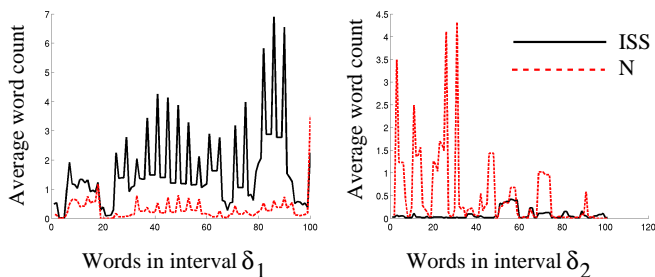

Fig. 8. Histogram of the word counts for two classes of skull shapes computed from the co-occurence matrix shown in Fig. 7.

III. RESULTS

A. BOW representation

We computed co-occurrence tables of counts for the BOW representation of the skull outlines in our population sample. Our data revealed that this symbolic representation encodes distinctive shape information that can be used to differentiate sagittal and normal skull shapes. Figure 8 shows a histogram of the word counts for two classes of skull shapes. The class of sagittals is represented as a black solid line and the class of normal skull shapes as a dashed red line. The histograms are computed in the intervals δ_1 and δ_2 from the co-occurence matrix shown in Fig. 7. It is clear from the figure that the distributions of word counts for each of the classes differ significantly.

B. Regression models

Our analysis shows that SSI and SSD descriptors are able to predict mental (MDI) and motor (PDI) scores, respectively. The error rates are reasonably low and highly significant, as can be seen in Table 1. It is worth noting that that the SSI shape descriptors reported in Table I are 2-D vectors whose components are individual SSI-A and SSI-F values. Shape descriptors that used SSI-M values had a poor performance (not reported in Table I).

	MDI		PDI	
	E	p-value	E	p-value
SSI	9.128	**0.0005**	8.582	0.5700
PR	9.373	0.0550	8.583	0.0380
CS	8.927	0.7040	8.584	0.0370
SSD	9.393	0.0190	8.467	**0.0000**

TABLE I

REGRESSION RESULTS BETWEEN SHAPE DESCRIPTORS (SSI, PR, CS, SSD) AND BEHAVIORAL SCORES (MDI AND PDI) . $N_V = 200$, $N_L = 3$ AND $N_R = 50$. KEY: $E = \sqrt{\text{AMSE}}$.

IV. DISCUSSION AND FUTURE WORKS

Even though it appears certain that isolated synostosis is associated with elevated risk for mild to moderate neurocognitive impairement, the causal mechanism that explains this association is still largely unknown. One possibility is that a) early suture fusion leads to a secondary brain deformation and, consequently, neuro-psychological deficits (Fig. 9a). Alternative notions propose that b) there is a yet unspecified primary neuropathology or mutation that leads to both synostosis and functional deficits, with no causal relation between skull abnormality and function (Fig 9b); and c) there are both primary and secondary processes that contribute to functional deficits [24], Figure 9c.
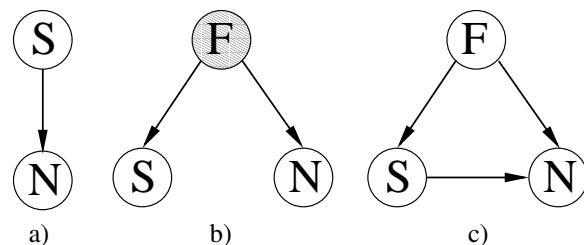

Fig. 9. Hypothetical causal mechanisms (represented as graphical models) relating synostitic skull shape (S), neurobehavioral dysfunction (N), and neuropathology or genetic mutation (F). a) Secondary brain deformation due to malformed skull shape. b) Primary neuropathology or mutation that leads to both synostosis and functional deficits, with no causal relation between skull abnormality and function (S is conditionally independent of N once F is observed; shaded node). c) There are both primary and secondary processes that contribute to functional deficits. Note that a graph without an arrow going from node F to node S is an alternative model for this hypothesis.

In this paper, we conducted regression analysis on skull shapes and neurodevelopmental condition. We found experimental evidence that the narrowing of the skull and the morphological variations captured by the SSDs, are

associated with an abnormal cognitive and psychomotor development, respectively. Although regression models only describe association among variables and do not imply a causal relationship [15], we suggest that hypotheses a) or c) are more likely to explain the observed neurophysiological dysfunctions.

It is clear from data in Table I that we need to refine our proposed regression models. This would allow us to improve error rates and verify their statistical significance. The sample size is also an important statistical variable. We are currently acquiring new data and hope to double the number of patients in our database. It is also desirable to test alternative nonlinear regression techniques such as Gaussian process [17], which enable the computation of the variance of the predicted values and minimize the possibility of over-fitting the data.

V. ACKNOWLEDGMENTS

REFERENCES

[1] Aldridge, K., Marsh, J. L., Govier, D., Richtsmeier, J. T. Central nervous system phenotypes in craniosynostosis. J of Anatomy **201**, 31-39. 2002.

[2] Arnaud, E., Renier, D., Marchac, D.: Prognosis for metnal function in scaphocephaly. Journal of Neurosurgery, **83**(3), 476-479, 1995.

[3] Bayley, N.: Manual for the Bayley Scales of Infant Development 2nd ed. San Antonio, TX: Psychological Corp., 1993.

[4] Fernback, S.K., Feinstein, K.A.: Radiologic evaluation of children with craniosynostosis. Neurosurgery Clinics of North America, **2**(3), 569-586, 1991

[5] Golland P., Fischl, B.: Permutation test for classifiers: towards statistical significance in image-based studies. In Proceedings of Information Processing in Medical Imaging (LNCS2732), 330-341, 2003.

[6] Guyon, I., Mikhoul, J., Schwartz, R., Vapnik, V.: What size test set gives good error rate estimates? IEEE Transactions on Pattern Analysis and Machine Intelligence **20**(1), 52-64, 1998.

[7] Haralick, R.M., Shapiro, L.G.: Computer and Robot Vision. Addison-Wesley, 1992.

[8] Hofmann, T. Unsupervised learning by Probabilistic Latent Semantic Analysis. Machine Learning **42**, 177-196, 2001.

[9] Kapp-Simon, K.A.: Mental development and learning disorders in children with single suture craniosynostosis. Cleft Palate-Craniofacial J. **35**(3), 197-203, 1998.

[10] Lajeunie, E., Le Merrer, M., Marchac, C., Renier, D.: Genetic study of scaphocepaly. Am. J. Med. Gene. **62**, 282-285, 1996.

[11] Lale, S.R., Richtsmeier, J.T.: An invariant approach to statistical analysis of shape. Chapman and Hall/CRC, 2001.

[12] Lin, H.J., Ruiz-Correa, S., Shapiro, L.G., Hing, A.V., Cunningham, M.L., Speltz, M.L., Sze, R.W.: Symbolic shape descriptors for classifying craniosynostosis deformations from skull imaging. In press, Proceedings of the IEEE Conference on Engineering in Medicine and Biology Society, 2005.

[13] Lin, H.J., Ruiz-Correa, S., Sze, R.W., Cunningham, M.L., Speltz, M.L., Hing, A.V., Shapiro, L.G.: Efficient symbolic signatures for classifying craniosynostosis skull deformities. Computer Vision for Biomedical Image Application, Lecture Notes in Computer Science, Springer-Verlag, Berlin, 302-313, 2005.

[14] Magge, S.N., Westerveld, M., Pruzinsky, T., Persing, J.A.: Long-term neuropsychological effect of sagittal craniosynostosis on child development. J. Craniofacial Surgery **13**(1), 99-104, 2002.

[15] Pearl, J.: Causality: Models, reasoning and inference. Cambridge University Press, 2000.

[16] Pershing, J., Magge, S.N., Westerveld, M., Pruzinsky, T.: Neuropsychological outcomes following sagittal craniosynostosis. Presented at the annual meeting of the American Cleft Palate-Craniofacial Association, Scottsdale, AZ, 1999.

[17] Rasmussen C., Williams, C. K. I. Gaussian Process for Machine Learning, The MIT Press, 2006.

[18] Rao, C.: Geometry of circular vectors and pattern recognition of shape of a boundary. Proc. Nat. Acad. Aci. **95** 12783, 2002.

[19] Renier, D., Lajeunie, E., Arnaud, E., Marchac, D.: Management of craniosynostosis. Child's Nerv Syst **16**, 645-658, 2000.

[20] Ruiz-Correa, S., Sze, R.W., Lin, H.J., Shapiro, L.G., Speltz, M.L., Cunningham, M.L.: Classifying craniosynostosis deformations by skull shape imaging. Proceedings of the IEEE Conference on Computer-Based Medical Systems, 335-340, 2005

[21] Ruiz-Correa, S., Sze, R.W., Starr, J.R., Hing, A.V., Lin, H.J., Cunningham, M.L.: A Fourier-based approach for quantifying sagittal synostosis head shape. American Cleft Palate-Craniofacial Association Meeting, 2005.

[22] Ruiz-Correa, S., Sze, R.W., Starr, J.R., Lin, H.J., Speltz, M.L., Cunningham, M.L., Hing, A.V.: New scaphocephalyseverity indices of sagittal craniosynostosis: A comparative study with cranial index quantifications. Cleft Palate-Craniofacial Journal **43** 2, 211-221, 2006.

[23] Scholköpf, B., Smola, A.J.: Learning with Kernels. Cambridge University Press, 2002.

[24] Speltz, M.L., Kapp-Simon, K.A., Marsh, J., Dawson, N.G.: Single suture synostosis: a review of neurobehavioral research and theory. Journal of Pediatric Phsychology, 29(8), 651-668, 2004.

[25] Virtanen, R., Korhonen, T., Fagerholm, J., Viljanto, J.: Neurocognitive sequelae of scaphocephaly. Pediatrics **3** 791-795, 1999.