3D Head Shape Quantification for Infants with and without Deformational Plagiocephaly

I. Atmosukarto[1], M.S., L. G. Shapiro[1], Ph.D., J. R. Starr[2], Ph.D., C. L. Heike[2], M.D, M.S., B. Collett[2], Ph.D., M. L. Cunningham[2], M.D, Ph.D., M. L. Speltz[2], Ph.D.

[1] Department of Computer Science and Engineering, University of Washington

[2] Seattle Children's Hospital

Corresponding author:

Matthew L. Speltz

Professor of Psychiatry & Behavioral Sciences,

University of Washington School of Medicine

Chief, Outpatient Psychiatry Services

Seattle Children's Hospital

Phone: 206-987-7577

Fax: 206-987-2246

Email: mspeltz@u.washington.edu

This publication was made possible in part by Grant Number 1 R01 HD046565 National Institute of Child Health and Human Development (NICHHD) to Dr. Speltz Grant Number 1 UL1 RR025014 from the National Center for Research Resources (NCRR), components of the National Institutes of Health (NIH)), and NSF Grant Number DBI-0543631 (PI: L. Shapiro). Dr. Heike was supported by Grant Number K23-DE017741. We also received support the General

Clinical Research Center (M01-RR00037).Its contents are solely the responsibility of the authors and do not necessarily represent the official view of NICHHD, NCRR, or NIH.

Running title: 3D Head Shape Quantification for Infants

Abstract

Objective: We developed and tested three dimensional (3-D) indices for quantifying severity of deformational plagiocephaly (DP).

*Design***:** We evaluated the extent to which infants with and without DP (as determined by clinic referral and two experts' ratings) could be correctly classified.

Participants: Infants ages 4-11 months, including 154 with diagnosed DP and 100 infants without a history of DP or other craniofacial condition. After excluding participants with discrepant expert ratings, data from 90 infants with DP and 50 infants without DP were retained.

Measurements: Two-dimensional histograms of surface normal vector angles were extracted from 3-D mesh data and used to compute the severity scores below.

Outcome measures: Left Posterior Flattening Score (LPFS), Right Posterior Flattening Score (RPFS), Asymmetry Score (AS), Absolute Asymmetry Score (AAS) and an approximation of a previously described 2-D measure, the Oblique Cranial Length Ratio (aOCLR). Two-dimensional histograms localized the posterior flatness for each participant.

Analysis: We fit receiver operating characteristic curves and calculated the area under the curves (AUC) to evaluate the relative accuracy of DP classification using the above measures.

Results: The AUC statistics were: AAS=91%; LPFS=97%, RPFS=91%; AS=99%, and aOCLR=79%.

*Conclusion***:** Novel 3-D-based plagiocephaly posterior severity scores provided better sensitivity and specificity in the discrimination of plagiocephalic and typical head shapes than the 2-D measurements provided by a close approximation of OCLR. These indices will allow for more precise quantification of the DP phenotype in future studies on the prevalence of this condition, which may lead to improved clinical care.

Keywords: plagiocephaly, head shape

INTRODUCTION

Deformational plagiocephaly (DP) refers to cranial asymmetry resulting from external forces shaping the infant's malleable skull such as positional preference or sleep positioning (McKinney et al., 2008). Common manifestations include a parallelogram shaped skull, with asymmetric flattening of the occiput, ipsilateral frontal bossing, and contralateral occipital bulging. Brachycephaly is thought to have similar etiology, and refers to bilateral (often symmetric) flattening of the occiput (Graham et al., 2005). Although considered a minor cosmetic condition by many clinicians, elevated rates of neurodevelopmental delay have been observed in some infants with DP and/or brachycephaly, which are especially evident in motor functions (Panchal et al., 2001; Kordestani et al., 2005; Speltz et al., 2009). DP is of concern to many parents because of its presumably negative effect on craniofacial appearance, although studies demonstrating such outcomes are limited (Bialocerkowski et al., 2005; Collett et al., 2005; Steinbok et al., 2007). Nevertheless, referrals to pediatricians, surgeons and craniofacial centers for evaluation of DP and/or brachycephaly have increased greatly over the past decade (Argenta et al., 1996; Mulliken et al., 1999), as has the number of parents opting for physical therapy interventions or orthotic helmet therapy to "normalize" their infant's head shape (Kane et al., 1996; Turk et al., 1996; Mulliken et al., 1999; Bialocerkowski et al., 2005; Collett et al., 2005). Despite this increase in clinical interest and parental concern, there has been relatively little research on the causes, course and outcomes of DP (Collett et al., 2005). This may, in part, reflect the challenge of reliably capturing the phenotype for epidemiologic studies.

The severity of DP varies from mild flattening to marked asymmetry along a spectrum that is difficult to quantify, with unclear separation of "abnormal" from "normal." This may partially account for the wide variation in reported prevalence estimates, which range from less than one percent to forty-eight percent (Hutchison et al., 2004, Bialocerkowski et al., 2008). Clinically, the assessment and diagnosis of DP is usually based upon a healthcare provider's physical examination of the infant's skull (and radiographic evaluation to rule out craniosynostosis, if suspected). This process is often subjective and may be biased by the diagnostician's knowledge of the infant's referral status (i.e., knowing that a parent or health provider is concerned about the infant's head shape).

Various methods for quantifying head shape have been developed for both clinical and research purposes (Hutchison et al., 2005; Ruiz-Correa et al., 2008) and several methods have been used to capture variation in DP. These include parental and clinician ratings and classifications, and measurements involving cross-diagonal transcranial lengths and the angles or ratios between or among them (Hutchison et al., 2005). Direct measurements can be taken on the infant's head by using measuring tapes or calipers, though some investigators have reported greater precision and reliability from measurements obtained on two-dimensional photographs taken from a standardized distance (Zonenshayn et al., 2004).

In a recent and well-designed application of the photographic method for measuring DP severity, Hutchison et al., 2005 developed a technique called *HeadsUp* that incorporates a top view digital photograph of infant heads fitted with an elastic head circumference band equipped with adjustable color markers to identify landmarks. The resulting photograph is then automatically analyzed to obtain a number of quantitative measurements for head shape

including head circumference, cephalic index (CI), oblique cranial ratio length (OCLR), and ear angles.

Although promising, the Hutchison method requires subjective decisions regarding the placement of midline and ear landmarks and the selection of the posterior point of the OCLR lines (40 degrees). In addition, as the authors noted (Hutchison et al., 2005), a head circumference band only represents head shape in two dimensions. As a result, displacement of head volume in the parietal and lower occipital areas cannot be assessed. Logistically, placing the band and capturing the photograph on an infant can also be challenging.

The goal of the present study was to develop and test a three-dimensional, automated procedure for the characterization of DP that does not depend on landmark selection. Using 3-dimensional (3-D) surface meshes of infants' skulls, we computed plagiocephaly posterior flattening scores with surface normal vectors; i.e., vectors that are perpendicular to the tangent plane at a particular point on a 3-D surface mesh. Based on the assumption that the surface normal vectors of 3-D points lying on the flat surfaces of the 3-D head meshes would have more similar angles than the surface normal vectors of 3-D points that lie on rounded surfaces, we hypothesized that these measures would differentiate plagiocephaly from more typical head shapes. We also performed exploratory analyses to determine whether these novel 3-D-based plagiocephaly posterior severity scores provide better sensitivity and specificity in the discrimination of plagiocephalic and typical head shapes than the 2-D measurements provided by OCLR.

METHODS

Participants

The study group initially comprised 254 participants: 154 infants referred to a craniofacial center and diagnosed with DP by a craniofacial specialist, and 100 infants without any diagnosed craniofacial anomalies who were recruited from a participant registry..

Cases. Cases were eligible if they were between the ages of 4 and 11 months at the time of diagnosis with DP. Infants were diagnosed with DP by one of four craniofacial specialists, all of whom were either pediatricians or Advanced Registered Nurse Practitioners. Exclusions included: (1) prematurity (less than 35 weeks gestation); (2) presence of a known neurodevelopmental condition, brain injury, or significant vision or hearing impairment; (3) major malformations or three or more minor malformations; (4) hemifacial microsomia; (5) a non-English speaking mother (6) a history of adoption or out-of-home placement; and (7) current plans for the family to move out of state before the completion of the project. Cases were seen for their initial study visit within 3-weeks of diagnosis on average (standard deviation = 1.0). These 154 enrolled cases represented 56% of all eligible cases. One hundred twenty one families declined to participate. When possible, we queried families about their reason(s) for choosing not to participate. Among 65 families providing one or more reasons for declining, 32 cited distance or transportation issues, 33 cited time constraints, 8 reported that they were unconcerned about their child's development, 1 cited privacy concerns, and 3 identified other reasons.

All participants were enrolled after obtaining informed consent approved by the Institutional Review Board of Seattle Children's Hospital. This research is in full compliance with HIPAA standards.

Controls. Infants were eligible for participation as "controls" if they (1) had no history of referral for or provider or parent concern about head shape or other craniofacial anomaly; and (2) did not meet any of the exclusionary criteria for cases described above. We identified nearly all

controls through an infant participant pool, consisting of families residing in King and Snohomish counties in Washington State who agreed at the time of their child's birth to be contacted for research participation at a later date. Families with a child in the target age range were contacted by phone to provide information about the study. Those who expressed an interest completed a brief phone screen to determine eligibility. The 100 enrolled control group participants represented 85% of all those who were screened by telephone. Fourteen infants were determined ineligible after phone screening and four declined participation. None of these four families provided a reason for choosing not to take part in the study.

Data Acquisition and Pre-processing

Each of the 254 participants was photographed by a medical photographer at the Seattle Children's Hospital. Each participant's head was first covered with a close-fitting cap to flatten the hair. Images of the participant were then taken using the 3-DMD cranial^{TM} imaging system (3-DMD, 2004), which uses four pods, each containing three cameras. Stereo analysis yields twelve range maps that are combined using 3-DMD proprietary software to produce a 3-D mesh of the subject's head. The resulting 3-D surface mesh data consists of 3-D point coordinates and the connectivity information among the points. We manually removed artifacts on images, such as clothing and noise; resulting in final 3-D head mesh data (Figure 1a).

All 3-D head mesh data were rotated and aligned in an automated manner in order to obtain the same pose and orientation across images (Wilamowska et al., 2009). Although faces are not completely symmetrical, pose alignment finds the yaw and roll angular rotations that minimize the difference between the left and right side of the face. The pitch of the head is then aligned by minimizing the difference between the height of the chin and the height of the

forehead. Less than 10% of the data required minor manual alignments to change the rotation parameter.

Expert Severity Ratings

The 3-D images of all cases and controls were assessed by two craniofacial dysmorphologists (authors MC and CH). Images were de-identified and viewed in random order to blind raters to case status. The experts created a scoring tool, which included graphical illustrations, and used the tool to assign discrete scores based on the severity of the posterior flattening by using a scale of 0 to 3 (i.e., 0 = normal, 1=mild, 2 = moderate, and 3= severe). They indicated the laterality of the flatness by using negative scores to represent left-sided flatness and positive scores to represent right-sided flatness (e.g., moderate left-sided flatness = -2; moderate right-sided flatness = 2).

Our objective was to develop and test a tool for assessing DP severity for which a "gold standard" does not exist. Hence, we considered the experts' ratings the gold standard in the evaluation of the new DP indices, yet the inter-rater agreement between the two dysmorphologists was only 65%. For the current analyses, we therefore excluded participants if (1) the two experts assigned discrepant posterior flattening scores (58 cases and 32 controls) or (2) the classification based on expert ratings differed from the clinical classification assigned at the time of enrollment (18 "controls" rated as having any posterior flattening, and 6 "cases" rated as having no DP by the experts). The final sample included 140 infants including 50 non-DP controls (by definition in category 0 by expert rating) and 90 cases, of whom 46 were in category -1 or 1, 35 in category -2 or 2, and 9 in category -3 or 3.

2-D Histogram of Azimuth-Elevation Angles of 3-D Surface Normal Vectors

Surface normal vectors were calculated for all points on the posterior side of the head. *Surface normal* vectors are perpendicular to the tangent plane at a particular point on a 3-D surface mesh. Surface normals can be computed as the vector cross-product of two non-parallel vectors on a surface. Given the surface normal vector $n(n_x, n_y, n_z)$ of a 3-D point, the azimuth angle θ of vector *n* is defined as the angle between the positive x-axis and the projection of vector *n* onto the x-plane. The elevation angle φ of *n* is defined as the angle between the x-plane and vector *n* (Figure 1b). The azimuth and elevation angles of surface normal vector *n* are calculated as follows:

$$\theta = \arctan(\frac{n_z}{n_x}) \quad \text{and} \quad \varphi = \arctan(\frac{n_y}{\sqrt{n_x^2 + n_z^2}})$$

where θ is in the range [-π, π] and φ is in the range [- π/2, π/2].

On relatively flat surfaces of the head, all surface normal vectors point in the same direction. By definition, individuals with DP should have one or more such flat areas, and larger such flat areas would have a higher number of surface normal vectors pointing in the same direction. Individuals without DP would be expected to have a more even distribution of the directions of surface normal vector, since such individuals' heads are more rounded (Figure 2). These observations provide the rationale for development of the four novel DP severity scores presented here.

After calculating the surface normal vectors of all points on the posterior side of the head, we constructed a 2-D histogram of the resulting azimuth and elevation angles (hereafter referred to simply as a "2-D histogram"). The elevation angles span 180°, ranging from -90 to 90°, while the azimuth angles span 360° degree ranging from -180 to 180°. We then grouped the computed angles into a small number of "bins," 12 bins for elevation and 12 for azimuth. These form the basis for constructing a 144-bin 2-dimensional histogram, each bin representing an azimuth-

elevation combination corresponding to a particular area on the head. The value of each bin is the percentage of surface normal vectors with a particular azimuth-elevation angle combination. Since the surface normal vectors of points that lie on a flat surface are almost parallel, they will have similar azimuth-elevation angles. Thus, flat parts of the head will tend to have high-valued bins or peaks in the 2-D histogram. In comparison, the surface normal vectors of points that lie on a rounded surface will have many different angles and hence would be distributed over multiple histogram bins.

Posterior Flattening Score

We defined a severity score for the left and right sides of the back of the head (posterior flattening) using selected bins of the 2-D histogram. The *Left Posterior Flatness Score* (LPFS) is the sum of the histogram bins that correspond to the combination of azimuth angles ranging from -90° to -30° and elevation angles ranging from -15° to 45°, while the *Right Posterior Flatness Score* (RPFS) is the sum of the histogram bins corresponding to the combination of azimuth angles ranging from -150° to -90° and elevation angles ranging from -15° to 45° (Figure 3).

Asymmetry Score

We developed an *Asymmetry Score* (AS) that represents the difference between the RPFS and the LPFS. The AS quantifies the degree of asymmetry and also indicates which side is flatter, with negative AS values indicating that the left side is flatter (LPFS > RPFS). The absolute value of the asymmetry score, *Absolute Asymmetry Score* (AAS), allows us to compare our measurements to the OCLR described in previous studies (Hutchison et al., 2005).

Oblique Cranial Ratio Length

In the *HeadsUp* technique (Hutchison et al., 2005), digital photographs of infants' heads were taken from the vertex perspective while infants wore an elastic head circumference band

with adjustable color markers to identify landmarks. The resulting photograph of head and band was then automatically analyzed to obtain quantitative measurements of head shape including: (1) Cephalic index (CI), and (2) OCLR. The OCLR is the ratio of the longer to the shorter cross-diagonal length and quantifies cranial asymmetry. Since we did not use a head circumference band, we computed an approximation of the OCLR (aOCLR) by taking a top view snapshot of the 3-D head mesh and measuring the cross-diagonal length of the head contour in the snapshot.

Severity Localization

We also used the 2-D histogram to indicate the specific location of any posterior flattening. This was done by identifying points at which the surface normal vectors' azimuth and elevation angles corresponded to the 16 relevant histogram bins used in the severity score computations. Points at which the azimuth-elevation angle combinations corresponded to one of these relevant bins were marked and subsequently displayed on a color map (Figure 4). High bin values are represented by warm colors (red, orange, yellow), while low bin values correspond to cool colors (blue, cyan, green). A representative non-DP control participant with an expert score of zero has all bins colored in cool colors, i.e. with no angle combination that is relatively more prevalent than any of the other combinations (Figure 4). In DP cases with right and left posterior flatness, the increasing prevalence of red, orange, and yellow indicates increasing severity of DP (Figure 4).

Data Analyses

We fit receiver operating characteristic (ROC) curves and calculated the area under the curves (AUC) to determine the extent to which DP cases and non-DP controls could be correctly classified by the LPFS, RPFS, AS, and aOCLR scores. For all possible diagnostic threshold values the ROC curve plots the sensitivity (percentage of cases correctly identified) versus one

minus the specificity (the percentage of non-DP head shapes correctly identified). To estimate overall accuracy, we computed the area under the ROC curve (AUC). A perfect diagnostic test yields an AUC of 1. We also selected a threshold value for each score, such that the threshold maximized the combination of sensitivity and specificity for distinguishing head shape characteristics such as left posterior flattening, right posterior flattening, and head asymmetry.

RESULTS

Descriptive Statistics

A higher proportion of cases than controls were male or of mixed race, and cases tended to be slightly younger and of lower socioeconomic status (Table 1). DP cases with left posterior flattening had a higher mean LPFS ranging from 0.159-0.194 (depending on the expert severity rating), while non-DP controls and DP cases with right posterior flattening had a lower mean LPFS ranging from 0.111-0.127 (Table 2 and Figure 5). In contrast, DP cases with right posterior flattening had mean RPFS ranging from 0.171-0.184 , while non-DP controls and DP cases with left posterior flattening had lower mean RPFS ranging from 0.115-0.144 (Table 3 and Figure 6). DP ases with left posterior flattening had mean AS ranging from -0.015 to -0.079, while DP cases with right posterior flattening had mean AS ranging from 0.048 to 0.069 (Table 4 and Figure 7). The non-DP control group had a slightly positive mean AS of 0.012. The distribution of the AAS for non-DP controls had a mean of 0.016 and standard deviation 0.012 (Table 5), while DP cases had higher mean AAS ranging from 0.042 – 0.073 (or 260-450% that of the non-DP control group mean). The mean aOCLR score for non-DP controls was 103.5 and ranges between 105.2-114.8 for DP cases (or 102-111% of the control group mean) depending on the assigned expert scores for these DP cases.

Graphical analyses indicated that there was no single threshold for any of the indices that perfectly distinguished DP cases and non-DP controls (Figures 5-9). Nevertheless, the automatically set threshold of 0.15 for the LPFS (or RPFS) distinguished most DP cases with left (or right) posterior flattening (enclosed in a box in Figures 5 and 6) from most non-DP controls and participants with right (or left) posterior flattening. Excluding the non-DP control participants, an AS threshold of zero produced a relatively clear distinction between DP cases with left and right posterior flattening (Figure 7). Setting the AAS threshold to 0.0352 provided a reasonable, though imperfect, classification of non-DP control participants (expert score = 0) versus DP case participants (expert score > 0) (Figure 8).

The AAS correlated with the aOCLR; for both measures, there was overlap in the range of scores between non-DP controls and DP cases who were given expert ratings of mild or moderate DP (Figure 9). All DP cases with expert ratings of severe DP were above the diagnostic threshold for both measures (0.035 for AAS and 106 for aOCLR).

Relative Discrimination of DP Cases and non-DP Controls

The LPFS and RPFS had relatively high accuracy in distinguishing DP cases from non-DP controls (Figures 10a and b), but they were not directly comparable with the aOCLR. The AAS produced more accurate classification (90.9%) than the aOCLR (78.6 %; Table 6 and Figure 10d). The AS demonstrated very high accuracy in the classification of DP cases with left posterior flattening from those with right posterior flattening (Figure 10c).

DISCUSSION

In this study we developed an automated procedure for the 3-D characterization of DP that does not depend on human landmark selection, a process that is subjective, time-consuming and

potentially unreliable. Three-dimensional representation is of special importance to the quantification of DP, as it allows for measurement of volume displacement in the parietal and lower occipital regions, areas that are not covered by 2-D methods. In addition to creating novel indices of flatness and asymmetry, we provided an initial test of their relative ability to discriminate DP cases and non-DP controls by comparing the new indices with a published 2-D method (OCLR; Hutchison et al, 2005). An important design feature of this study was the operational definition of DP on the basis of both: (1) the occurrence of clinic referral and (2) the independent, blinded ratings of 3-D images by two craniofacial dysmorphologists.

All four novel severity indices yielded high accuracy in the classification of DP cases vs. non-DP control group participants, with all classification coefficients greater than 90%. The asymmetry score (AS) provided the best overall discrimination (99.5%). In contrast, the aOCLR demonstrated lower accuracy of classification (78.6%), suggesting that the methods described here--surface normal vectors and resulting 2-D histograms--offer better definition of cranial asymmetry than the computation of cross-diagonal length ratios.

As expected, mean non-DP control group scores for all of the severity indices were lower than DP cases' average scores. However, the variance for non-DP control group scores was relatively high, as was variation among cases rated by the experts as having mild DP. In addition, we excluded DP cases and non-DP controls for which the expert scores disagreed, many of whom were in the zero-one range. For both reasons, these indices therefore may be less useful for distinguishing mild DP cases from individuals without DP.

Our goal in creating these severity indices was to develop a 3-D method for more precise, reliable, informative, and efficient quantification of variation in DP head shape. The new indices provide continuous (vs. categorical) measurement of head shape that could be used, for example,

to examine associations between DP severity and neurodevelopmental outcomes (e.g., see Speltz et al., 2009) or to assess changes in head shape over time, both for cases who undergo helmet therapy and those who do not. These procedures may also be applied in population-based studies to determine what constitutes normal variation in infant head shape and prevalence of DP. The new indices provide more information than previous methods, particularly the localization of the deformation, which might prove in future studies to have predictive significance. For example, the side of occipital flatness may have implications for the development of particular motor functions. These methods may also prove useful for the assessment of head shape in craniosynostosis, craniofacial microsomia, and other craniofacial conditions.

Clinical diagnosis and treatment may also benefit from the more precise quantification and localization offered by these novel indices, particularly if subtle distinctions in flatness, asymmetry, or location are associated with clinical outcomes. Diagnosticians may focus on the most severe flat areas of the head. However, a head shape with moderate left posterior flatness may have additional shape deformation in other areas of the skull that are not captured by the clinician's diagnosis, but nevertheless has potential diagnostic value (e.g., response to helmet therapy). This possibility requires further study. The potential clinical utility of our methods will depend on the outcomes of future research studies, as well as the availability and affordability of 3-D imaging systems and user-friendly software that automatically cleans images and computes the severity indices described here.

This study has several limitations. First, the comparison of the new indices with OCLR to discriminate DP cases and non-DP controls was limited by the fact that our participants were not marked with the necessary landmarks used in the HeadsUp program. Our computed aOCLR measurements relied upon approximations of the true OCLR indices (i.e., the cross-diagonal line

346 at 45 degrees from the vertical midline). Second, we did not include additional dysmorphic
347 features associated with DP that may have diagnostic or prognostic value (e.g., asymmetric ear
348 placement, forehead shape, and brachycephaly). We are working on other indices to capture such
349 features. Third, the two experts had relatively low inter-rater agreement in their ratings of
350 posterior flattening, which reflects the inherent difficulty of assigning categorical values to three-
351 dimensional shape, and may also reflect the fact that judgments were based on computer images,
352 not the infants themselves. A large number of cases and controls were excluded on the basis of
353 these ratings (due in part to the expected presence of mild plagiocephaly among infants enrolled
354 in the control group); this produced two highly contrasting groups of DP cases and non-DP
355 controls. Although the resulting isolation of two highly contrasting groups of DP cases and non-
356 DP controls suited the objectives of this study, the accuracy estimates may not generalize to a
357 more heterogeneous population with milder forms of posterior flattening. Still, it is possible that
358 the proposed methods would perform more accurately than existing methods even in such a more
359 heterogeneous group of children. Finally, we did not examine the correspondence between our
360 measures and other measures of brachycephaly (e.g., the cephalic index (CI) used by Hutchison
361 et al., 2005), though these comparisons are planned.

362 **CONCLUSION**

363 In this study we presented a new methodology to quantify and localize skull asymmetry
364 associated with DP. Two-dimensional histograms of surface normal vector angles were extracted
365 from the 3-D head mesh data from each participant in the study. Four severity scores, the LPFS,
366 RPFS, AS, and AAS, were computed from 2-D histogram representations, which can also be
367 used to localize and highlight the flat posterior areas. In this study, the proposed AAS more
368 accurately distinguished DP cases from non-DP controls than our approximation to the

previously described OCLR (aOCLR), which is based on measurements taken from 2-D photographs. Although our primary goal was to develop improved severity indices for research purposes, clinical diagnosis may also eventually benefit from their use, pending further research on the predictive value of variation in flatness and asymmetry, and location of dysmorphology.

REFERENCES

3-DMD. 3-DMDcranialTM System, Atlanta, GA. 2004.

Argenta LC, David LR, Wilson J, Bell WO. An increase in infant cranial deformity with supine sleep positioning. *J Craniofac Surg.* 1996; *7*: 5-11.

Bialocerkowski AE, Vladusic SL, Howell SM. Conservative interventions for positional plagiocephaly: a systematic review. *Developmental Medicine & Child Neurology*. 2005; 47:563-570.

Bialocerkowski AE, Vladusic SL, Ng CW. Prevalence, risk factors, and natural history of positional plagiocephaly: A systematic review. *Developmental Medicine & Child Neurolog.,* 2008; 50:577-586.

Collett B, Breiger D, King D, Cunningham M, Speltz M.L. Neurodevelopmental implications of "deformational" plagiocephaly. *Developmental and Behavioral Pediatrics*. Oct 2005; 26(5):379-389.

Graham JM, Kreutzman J, Earl D, Halberg A, Samayoa C, Guo X. Deformational brachycephaly in supine-sleeping infants.*J. Pediat.*2005; 146(2):253-257.

Hutchison BL, Hutchison LA, Thompson JM, Mitchell EA. Plagiopcephaly and Brachycephaly in the First Two Years of Life: A Prospective Cohort Study. *J. Pediat.* 2004, 114(4): 970-980.

392 Hutchison BL, Hutchison LA, Thompson JM, Mitchell EA. Quantification of plagiocephaly and
393 brachycephaly in infants using a digital photographic technique. *Cleft Palate Craniofac J*.
394 2005; 42(5):539-547.
395 Kane AA, Mitchell LE, Craven KP, Marsh JL. Observations on a recent increase in
396 plagiocephaly without synostosis. *Pediatrics.* 1996; 97: 877-885.
397 Kordestani RK, Patel S, Bard DE, Gurwitch R, Panchal J. Neurodevelopmental delays in
398 children with deformational plagiocephaly. *Plast Reconstruc Surg*. 2005; 117: 207-218.
399 McKinney CM, Cunningham ML, Holt VL, Leroux B, Starr J. Characteristics of 2733 cases
400 diagnosed with deformational plagiocephaly and changes in risk factor over time. *CleftPalate*
401 *Craniofac. J.* 2008;45(2):208-216.
402 Mulliken JB, Van Der Woude DL, Hansen M, LaBrie RA, Scott RM.. Analysis of posterior
403 plagiocephaly: Deformational versus synostotic. *Plast Reconstr Surg*. 1999; 103: 371-380.
404 Panchal J, Amirsheybani H, Gurwitch R, Cook V, Francel P, Neas B, Levine N.
405 Neurodevelopment in children with single-suture craniosynostosis and plagiocephaly without
406 synostosis. *Plast Reconstr Surg*. 2001; 108: 1492-1500.
407 Ruiz-Correa S, Starr JR, Lin HJ, Kapp-Simon KA, Sze RW, Ellenbogen RG, Speltz ML,
408 Cunningham ML. New severity indices for quantifying single-suture metopic
409 craniosynostosis. *Neurosurgery*. 2008; 63(2):318-324.
410 Speltz ML, Collett BR, Stott-Miller M, Starr JR, Heike C, Wolfram-Aduan A, King D,
411 Cunningham ML. Case-Control Study of Neurodevelopment in Deformational
412 Plagiocephaly. Unpublished manuscript (under editorial review), March, 2009.
413 Steinbok P, Lam D, Singh S, Mortensen PA, Singhal A. Long-term outcome of infants with
414 positional occipital plagiocephaly. *Childs Nerv Syst.*, 2007; 23(11):1275-1283.

Turk AE, McCarthy JG, Thorne CH, Wissof JH. The "back to sleep campaign" and DP: is there cause for concern? *J Craniofac Surg*, 1996; 7: 12-18.

Wilamowska K, Shapiro LG, Heike CL. Classification of 3-D face shape in 22q11.2 deletion syndrome. *IEEE International Symposium on Biomedical Imaging*, 2009.

Zonenshayn M, Kronberg E, Souweidane MM. Cranial index of symmetry: an objective semiautomated measure of plagiocephaly. Technical note. *J Neurosurg*. 2004 May;100(5 Suppl Pediatrics):537-540.

LIST OF FIGURES

Figure 1 (a) 3-D head mesh and orientation of a 3-D head mesh with respect to the x, y, and z axis in 3-D space (b) Azimuth angle θ and elevation angles φ of a surface normal vector n.

Figure 2 (a) The surface normal vectors of points that lie on a flat surface tend to have similar azimuth and elevation angles. (b) The surface normal vectors of points that lie on a more rounded surface have a wider distribution of angles.

Figure 3 (a) 2-D Histogram of azimuth elevation angles of surface normal vector angles. The Left Posterior Flattening Score is computed by summing the values of the bins highlighted in red, while Right Posterior Flattening Score is computed by summing the values of the bins highlighted in green. (b) Back view of the head showing the points whose surface normal vector angles correspond to the selected bins' azimuth-elevation angle combination highlighted in the 2-D histogram.

Figure 4(a): Mesh surface depictions of seven skulls representative of possible deformational plagiocephaly severity scores from expert clinician ratings. (b) Relevant bins of 2-D histogram of azimuth and elevation angles of surface normal vectors on 3-D head mesh models. These bins are used to calculate the various deformation severity indices. As the severity of posterior flatness increases on the side of the head, the peak in the 2-D histogram becomes more prominent as shown by the warmer colors (red, yellow, green). (c)The last row shows the localization of the posterior flatness, where the flat areas are colored in a similar shade as their corresponding histogram bins.

Figure 5 Correlations between Left Posterior Flattening Score and Expert Score. The optimal threshold at 0.15 (thick line) distinguishes the cases with left posterior flattening (enclosed in box) from the rest of the participants.

Figure 6 Correlations between Right Posterior Flattening Score and Expert Score. The optimal threshold at 0.15 (thick line) distinguishes the cases with right posterior flattening (enclosed in box) from the rest of the participants.

Figure 7 Correlations between Asymmetry Score and Expert Score. A threshold at value = 0 produces a clear distinction between cases with left posterior flattening (enclosed in the box in the lower left quadrant) and cases with right posterior flattening(enclosed in the box in the upper right quadrant).

Figure 8 Correlations between the Absolute Asymmetry Score and Expert Score. Setting threshold at value 0.0352 (thick line) provides a reasonable classification of non-DP control participants versus DP cases participants.

Figure 9 Correlation between the Absolute Asymmetry Score and the approximate Oblique Cranial Ratio Length.

Figure 10 (a) Receiver Operating Characteristic (ROC) using Left Posterior Flattening Score (LPFS) curves for classification of cases with left posterior flattening versus other participants.

The sensitivity and specificity at which the AUC is maximized (marked point on the graph) are 96.6% and 95.8% respectively.

Figure 10 (b) Receiver Operating Characteristic (ROC) curves using Right Posterior Flattening Score (RPFS) for classification of cases with right posterior flattening versus other participants. The sensitivity and specificity at which the AUC is maximized (marked point on the graph) are 91.9% and 86.4% respectively.

Figure 10(c) Receiver Operating Characteristic (ROC) curves using Asymmetry Score (AS) for classification of patients with left posterior flattening versus patients with right posterior flattening. The sensitivity and specificity at which the AUC is maximized (marked point on the graph) are 100% and 98.5% respectively.

Figure 10(d) Receiver Operating Characteristic (ROC) curves for classification of patients with posterior flattening versus non-DP controls using Absolute Asymmetry Score (AAS) and approximate Oblique Cranial Length Ratio (aOCLR). The performance of AAS is better than that of aOCLR. The sensitivity and specificity at which the AUC is maximized (marked point on the graph) are 96% and 80% respectively.

491

492 **Table 1 Demographic characteristics by group.**

493

Characteristic	DP Cases (n=154) n (%)	Non-DPControls (n=100) n (%)
Sex		
Male	102 (66.2)	61 (61.0)
Female	52 (33.8)	39 (39.0)
Age (months)		
4-5	49 (31.82)	42 (42.0)
6-7	71 (46.1)	32 (32.0)
8-9	23 (14.9)	22 (22.0)
10-11	11 8 (7.1)	4 (4.0)
Race/Ethnicity		
Caucasian	108 (70.1)	78 (78.0)
Asian/Pacific Islander	7 (4.6)	4 (4.0)
Black/African American	0 (0.0)	1 (1.0)
Latino	3 (2.0)	0 (0.0)
Native American	1 (0.7)	0 (0.0)
More than One Race	35 (22.7)	17 (17.0)
Familial SES		

I (High)	46	(29.9)	37	(37.0)
II	65	(42.2)	50	(50.0)
III	18	(11.7)	10	(10.0)
IV	11	(7.1)	3	(3.0)
V (Low)	2	(1.3)	0	(0.0)
Missing	12	(7.8)	0	(0.0)

494

495

Table 2 Descriptive statistics for the Left Posterior Flattening Score (LPFS)

Patient group	Expert Score	Mean	Standard Deviation
Non-DP control	0	0.127	0.014
DP cases with left posterior flattening	-1	0.159	0.018
	-2	0.182	0.025
	-3	0.194	0.040
DP cases with right posterior flattening	1	0.123	0.013
	2	0.116	0.014
	3	0.111	0.008

Table 3 Descriptive statistics for the Right Posterior Flattening Score (RPFS)

Patient group	Expert Score	Mean	Standard Deviation
Non-DP control	0	0.139	0.018
DP cases with left posterior flattening	-1	0.144	0.017
	-2	0.127	0.015
	-3	0.115	0.016
DP cases with right posterior flattening	1	0.171	0.023
	2	0.184	0.020
	3	0.181	0.020

Table 4 Descriptive statistics for the Asymmetry Score (AS)

Patient group	Expert Score	Mean	Standard Deviation
Non-DP control	0	0.012	0.016
DP cases with left posterior flattening	-1	-0.015	0.018
	-2	-0.055	0.020
	-3	-0.079	0.026
DP cases with right posterior flattening	1	0.048	0.024
	2	0.068	0.022
	3	0.069	0.015

Table 5 Descriptive statistics for the Absolute Asymmetry Score (AAS) and approximate Oblique Cranial Length Ratio (aOCLR) measurements

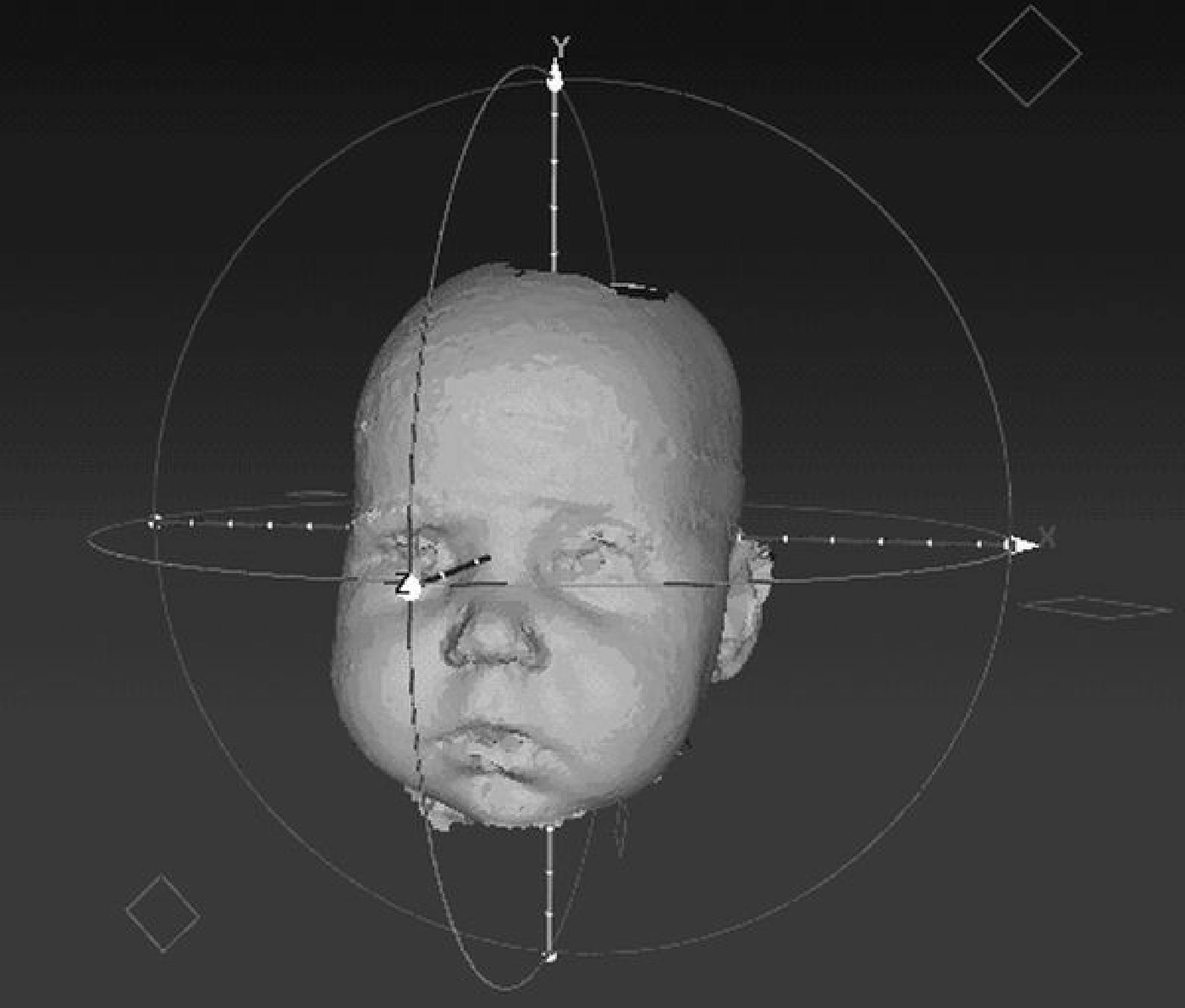
Patient group	Absolute Expert Score	Method	Mean	Standard Deviation
Non-DP control	0	AAS[+]	0.016	0.012
		aOCLR[+]	103.566	2.474
DP cases	1	AAS	0.042	0.024
		aOCLR	105.218	3.259
	2	AAS	0.064	0.022
		aOCLR	109.135	3.298
	3	AAS	0.073	0.020
		aOCLR	114.809	3.427

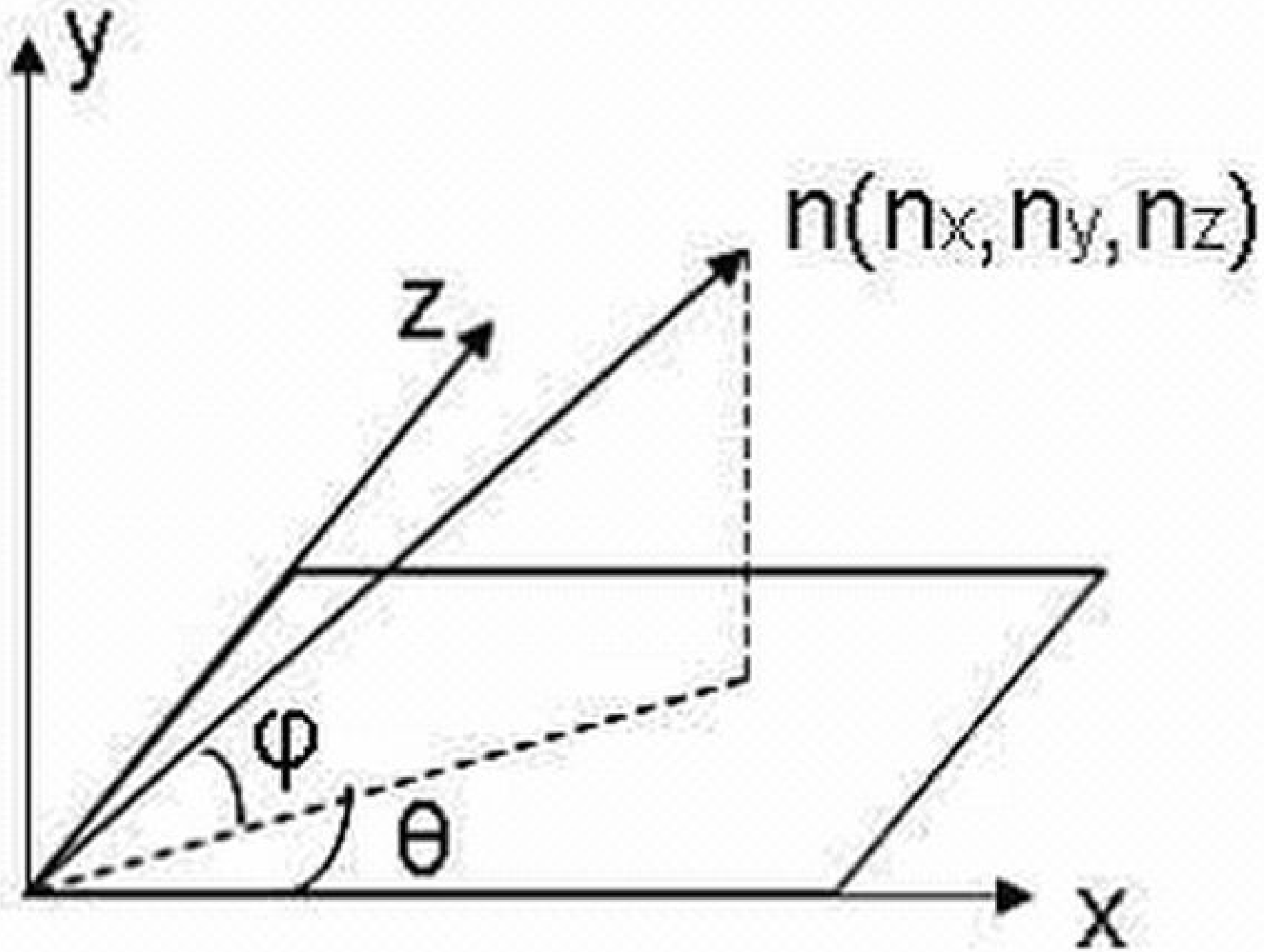
[+] AAS – Absolute Asymmetry Score, aOCLR – approximate Oblique Cranial Length Ratio

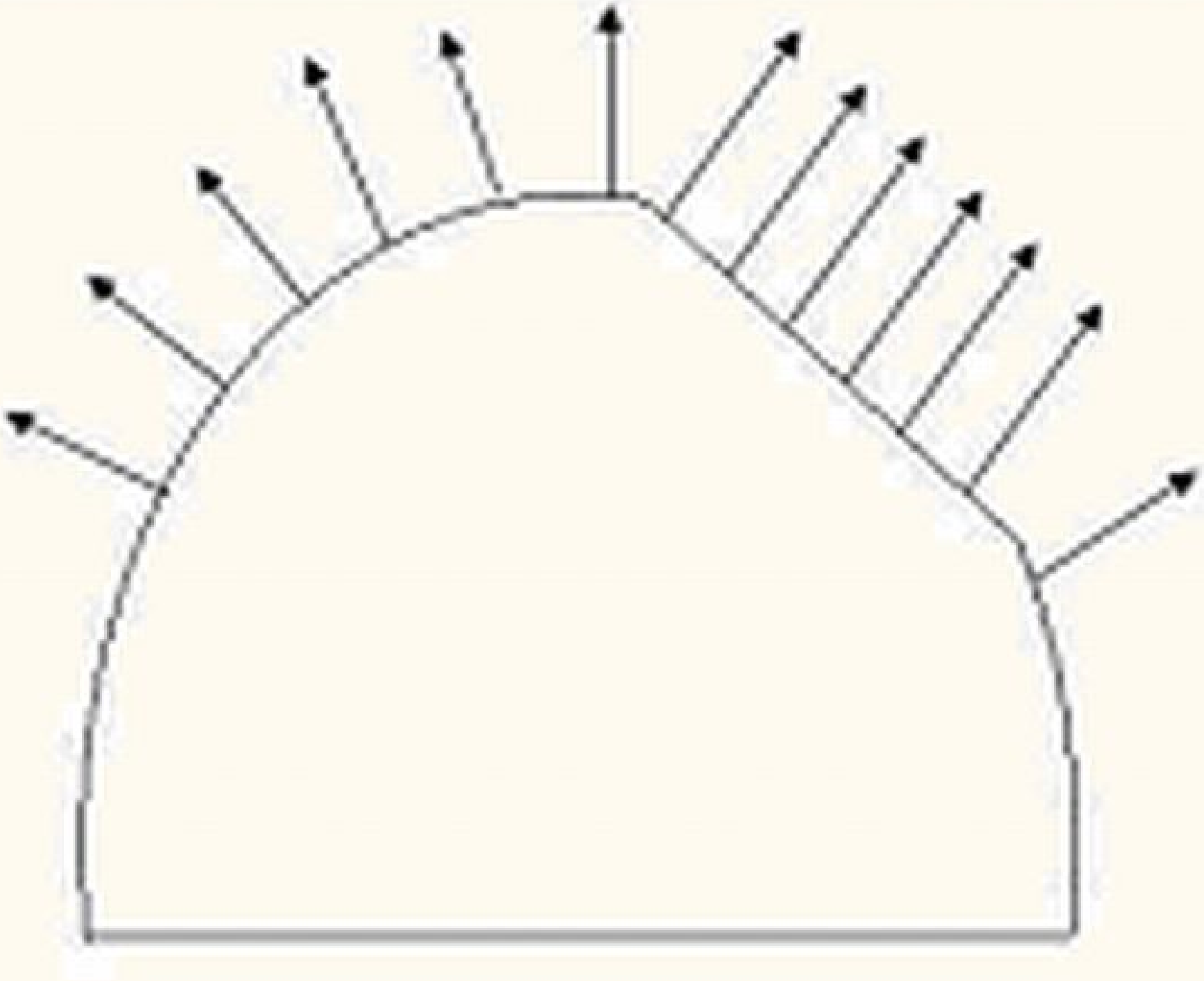
Table 6 Area under the curve and corresponding 95% confidence intervals computed from receiver operating characteristics curves

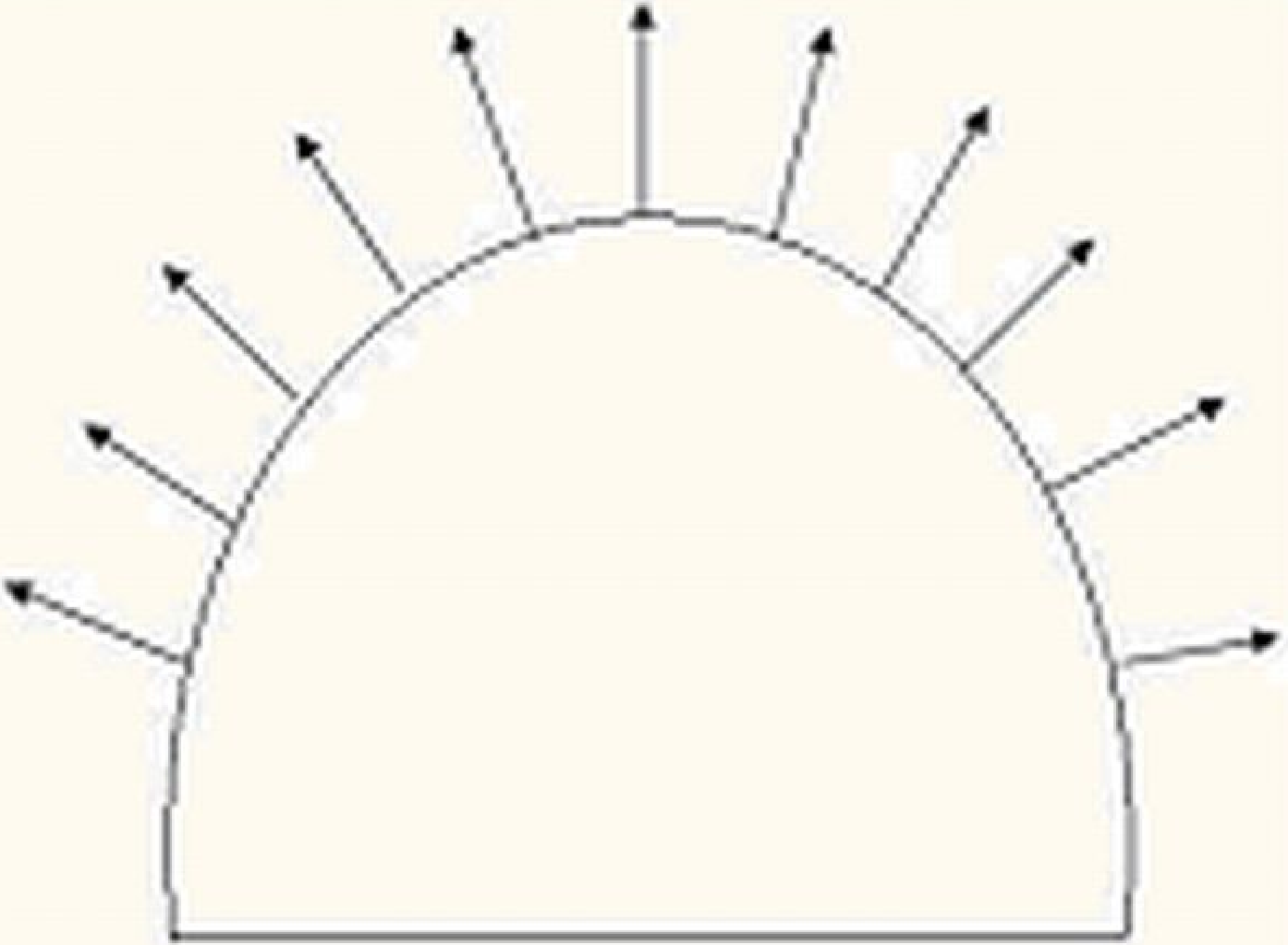
Score	AUC	95% CI
LPFS[+]	0.9745	0.93 – 1.02
RPFS[+]	0.91851	0.87 – 0.97
AAS[+]	0.90911	0.86 – 0.96
aOCLR[+]	0.78611	0.71 – 0.86
AS[+]	0.99558	0.98 – 1.01

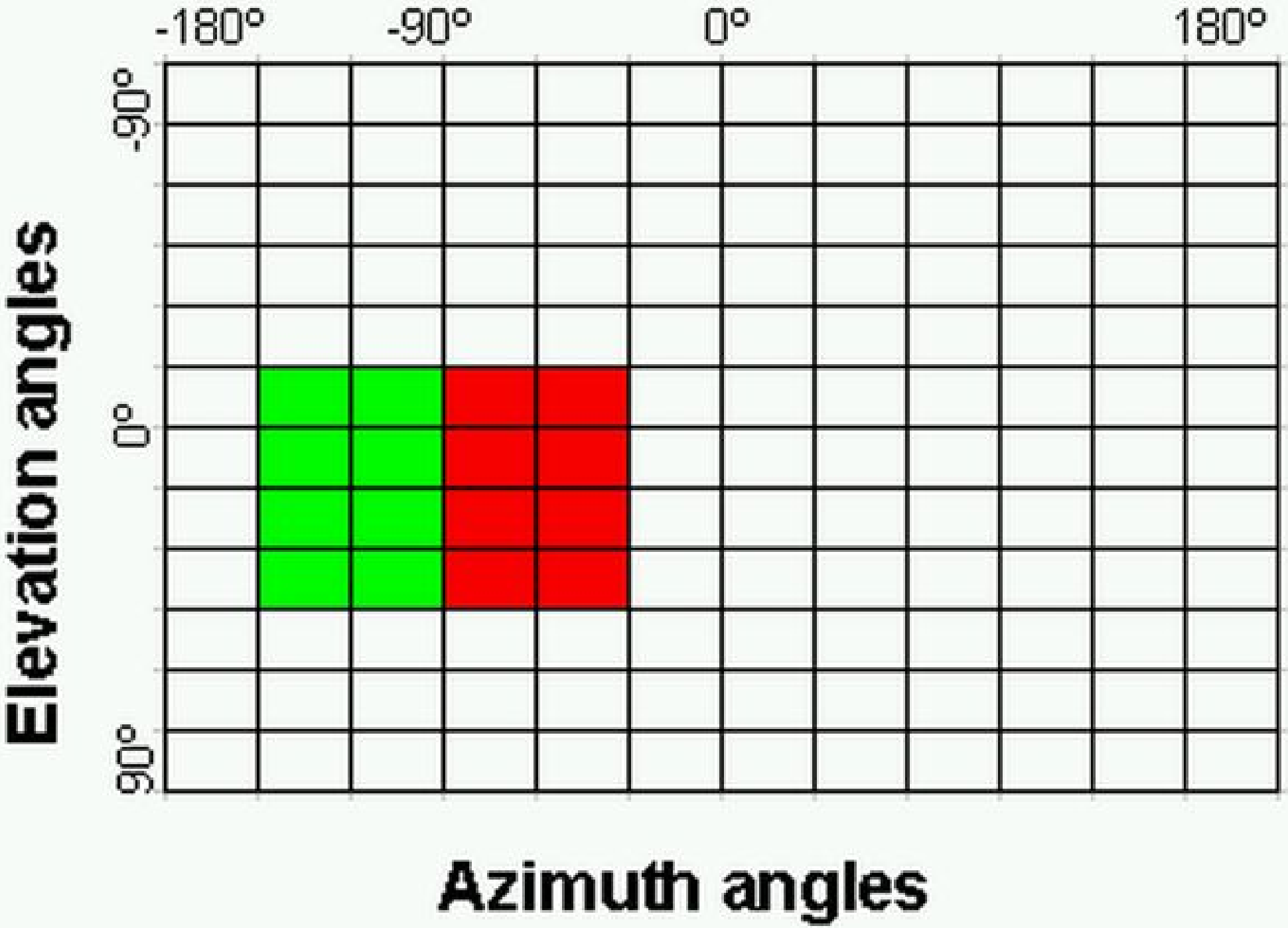
[+] LPFS – Left Posterior Flattening Score, RPFS – Right Posterior Flattening Score, AAS – Absolute Asymmetry Score, aOCLR – approximate Oblique Cranial Length Ratio, AS – Asymmetry Score

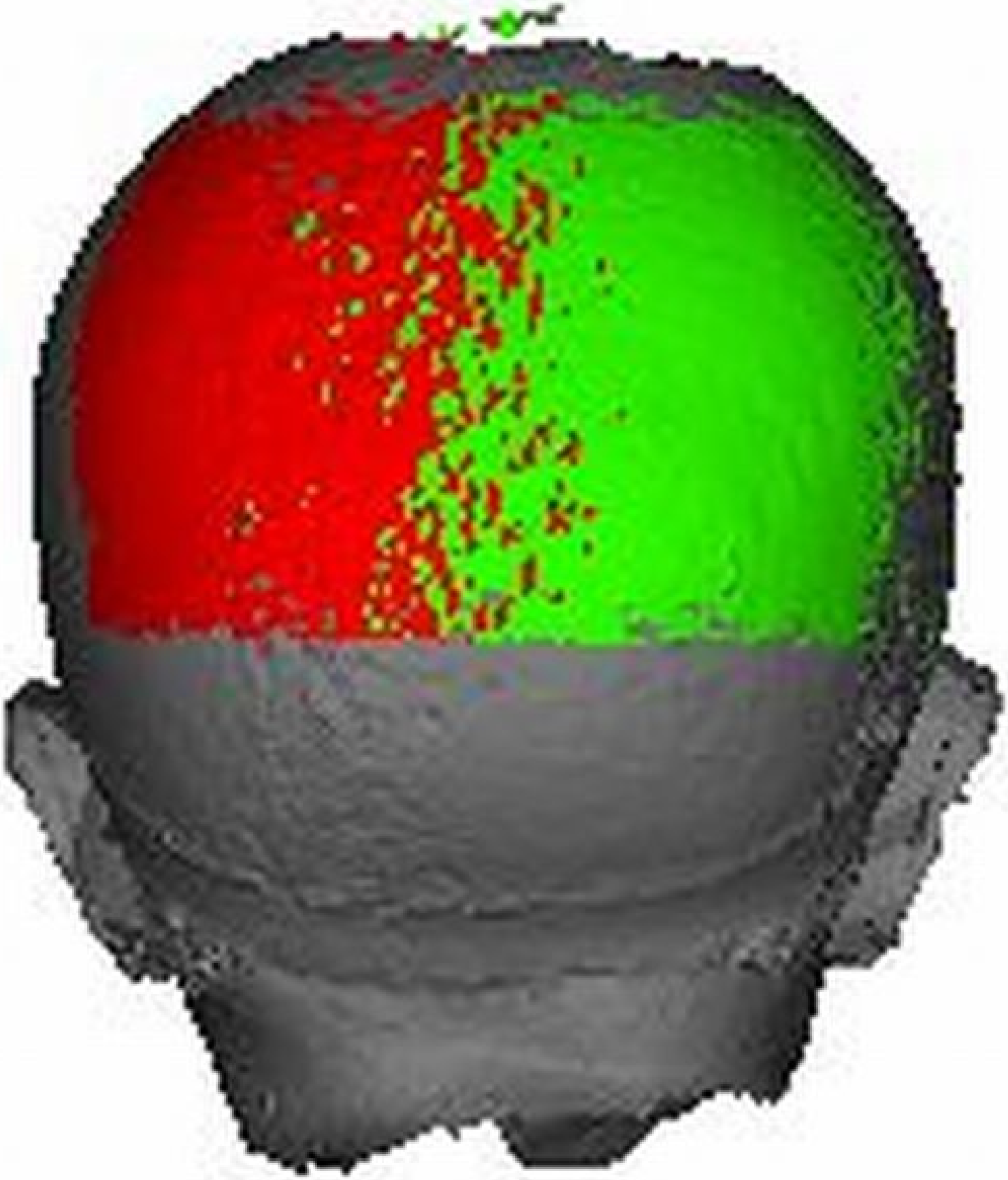

Y
X
Z

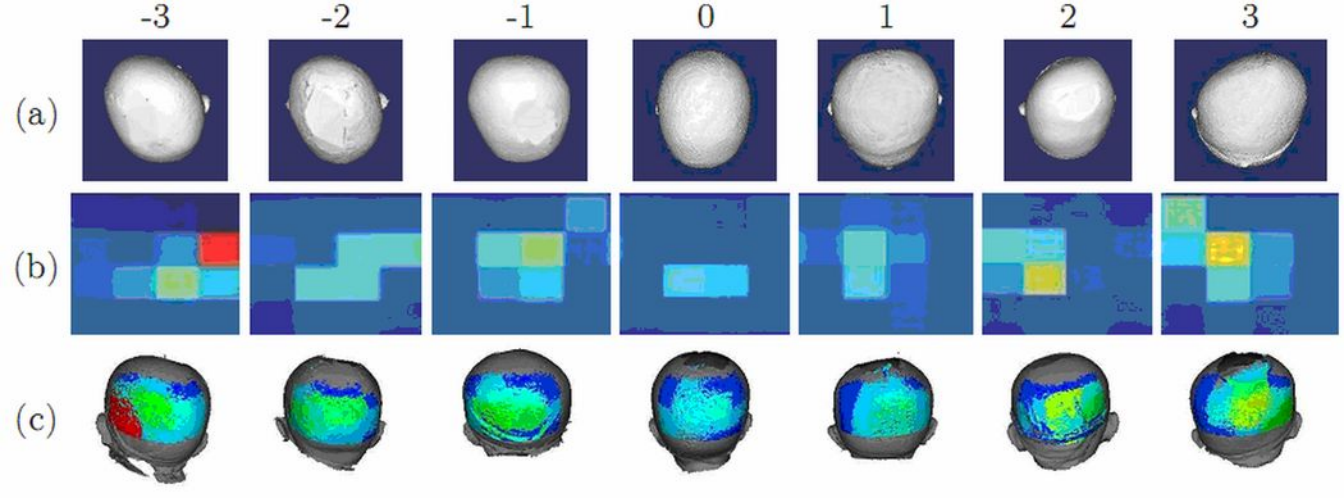

y
n(nx, ny, nz)
z
φ
θ
x

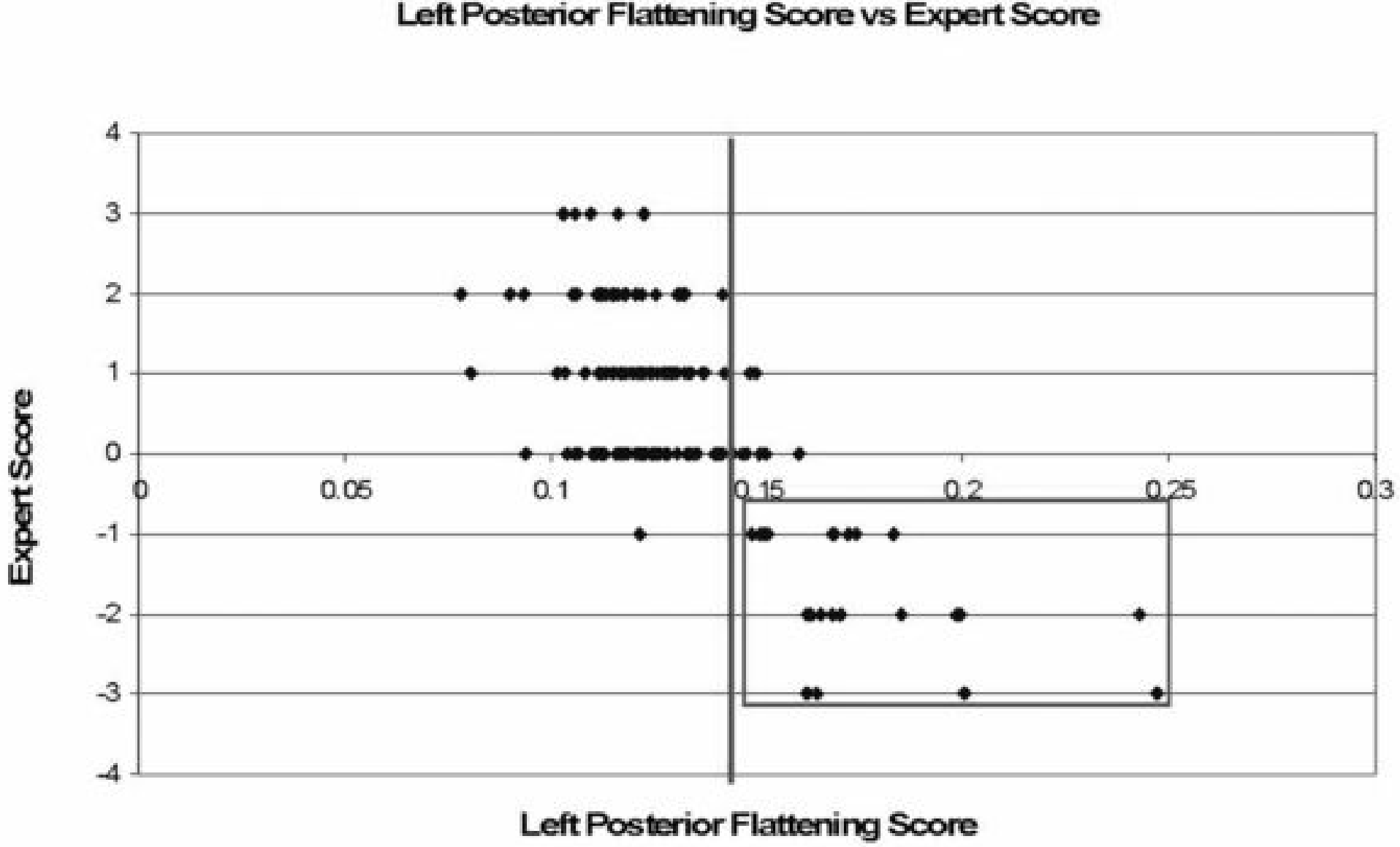


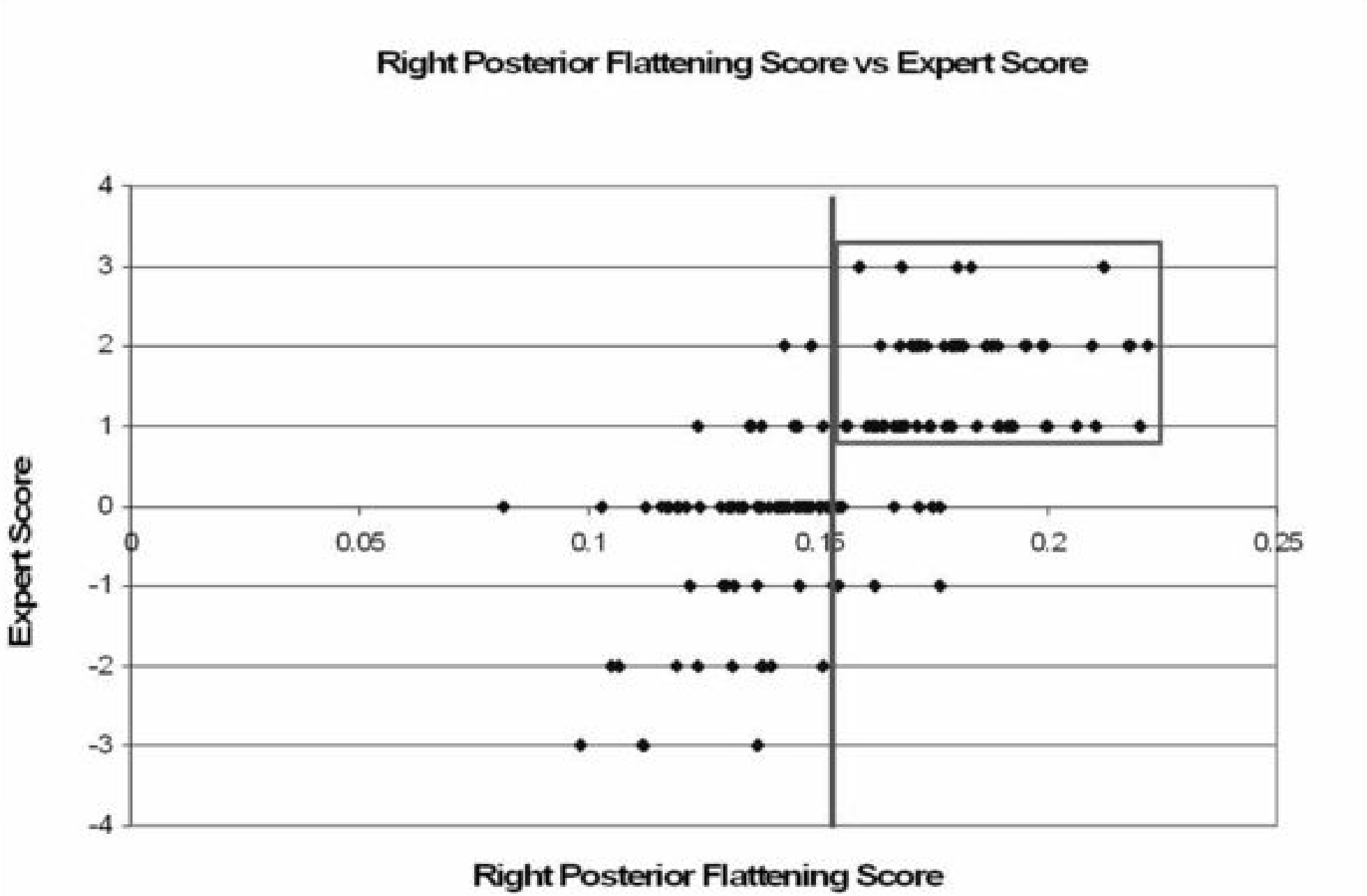


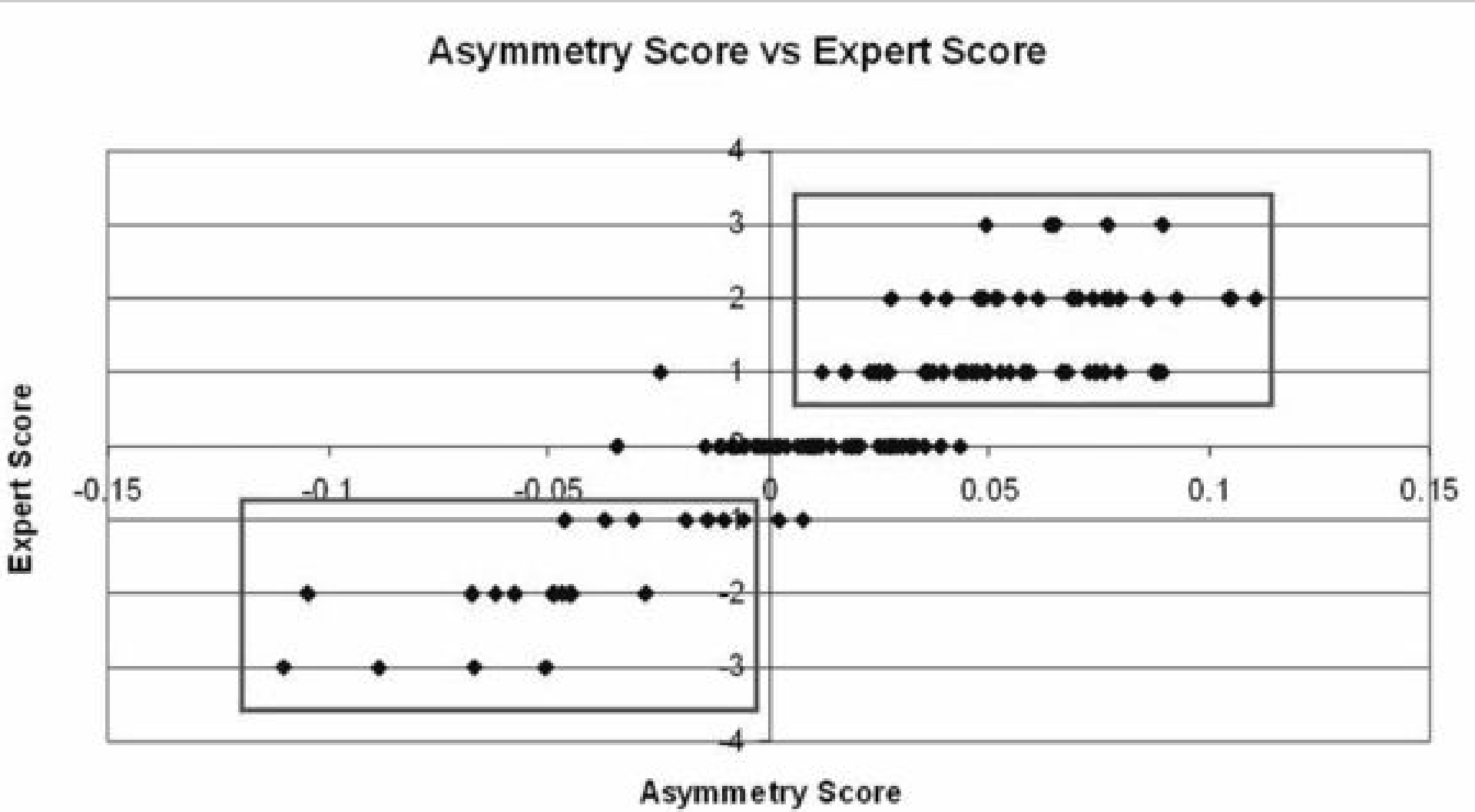

-180°
-90°
0°
180°
-90°
0°
90°
Elevation angles
Azimuth angles

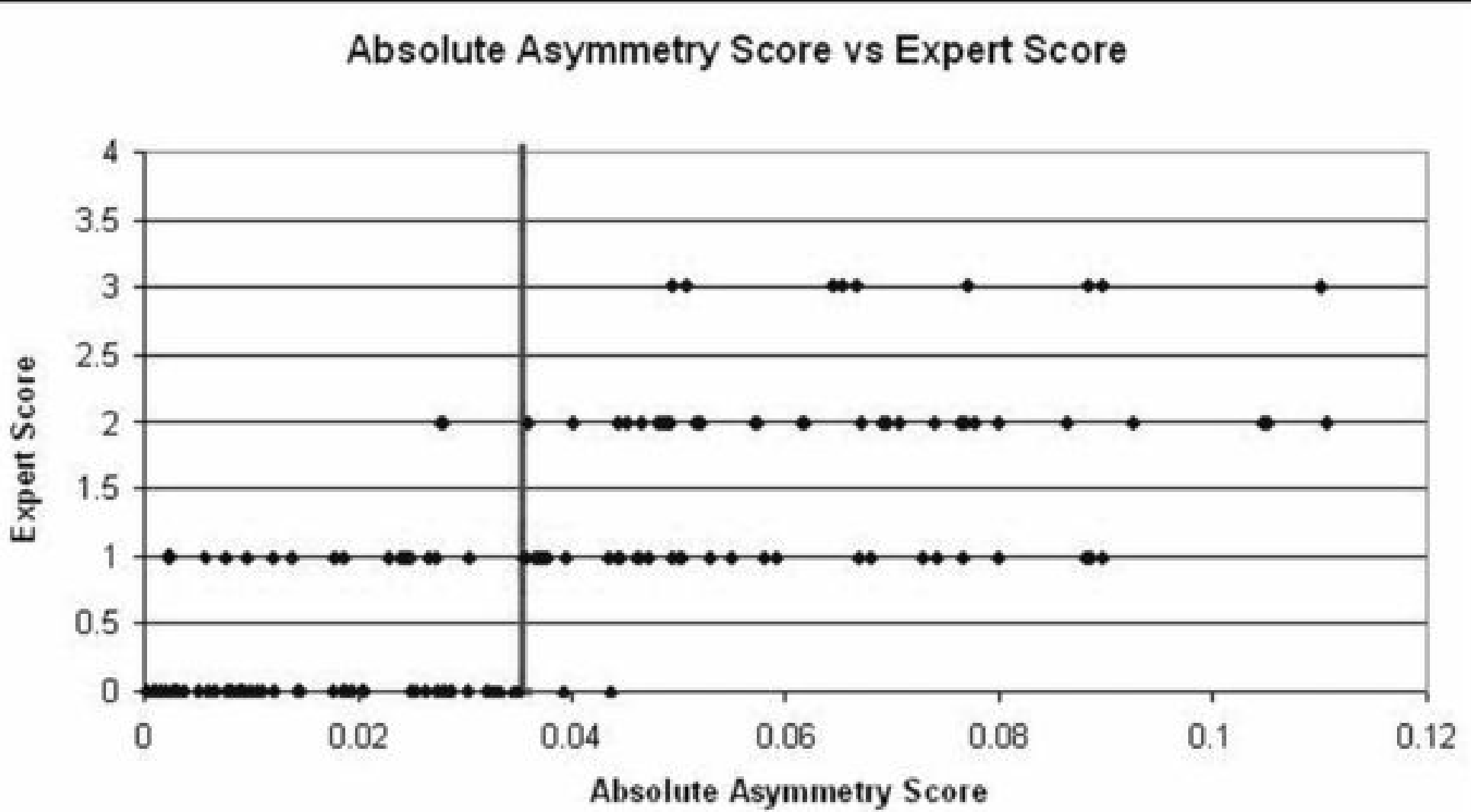


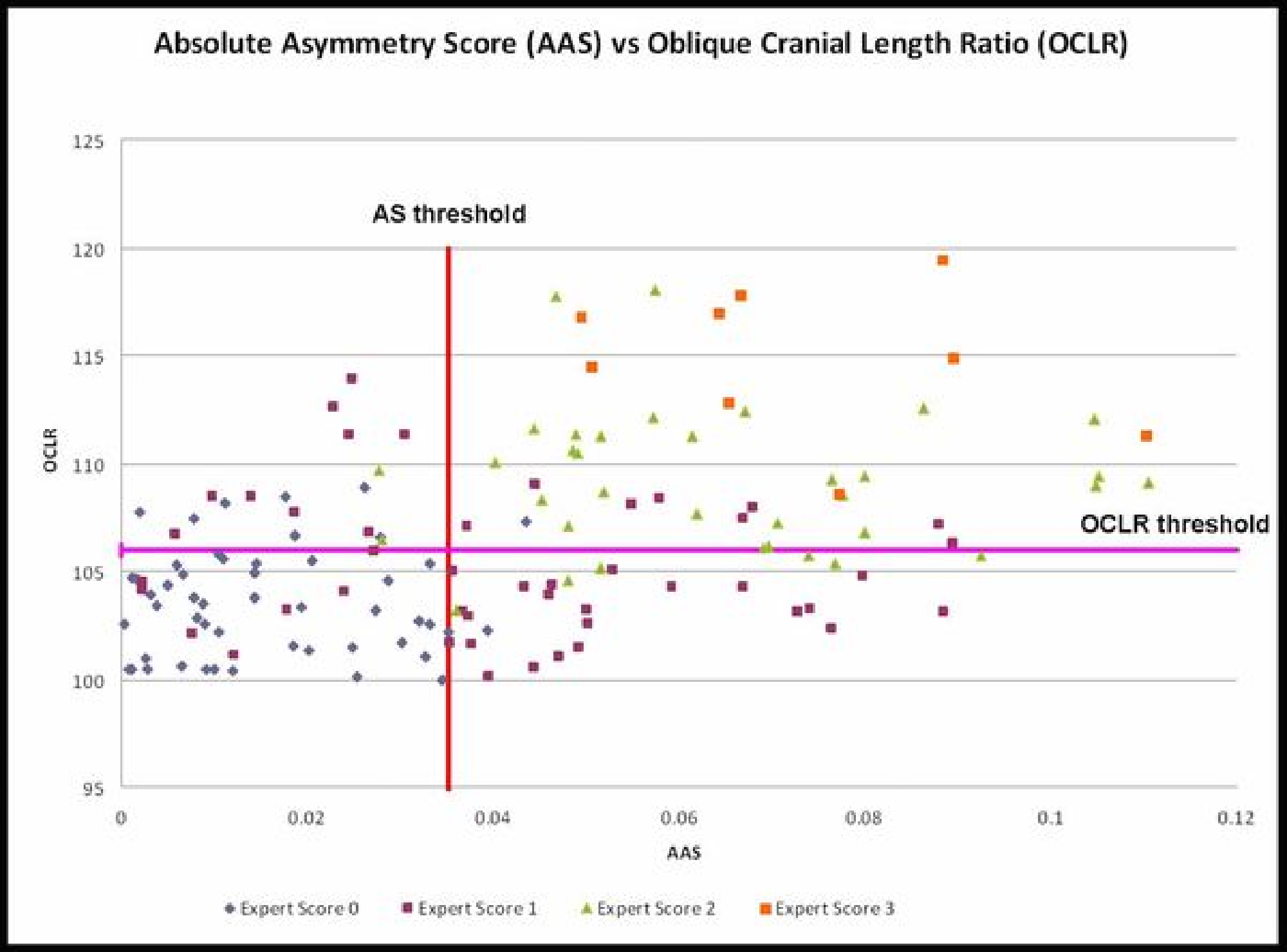

-3
-2
-1
0
1
2
3
(a)
(b)
(c)

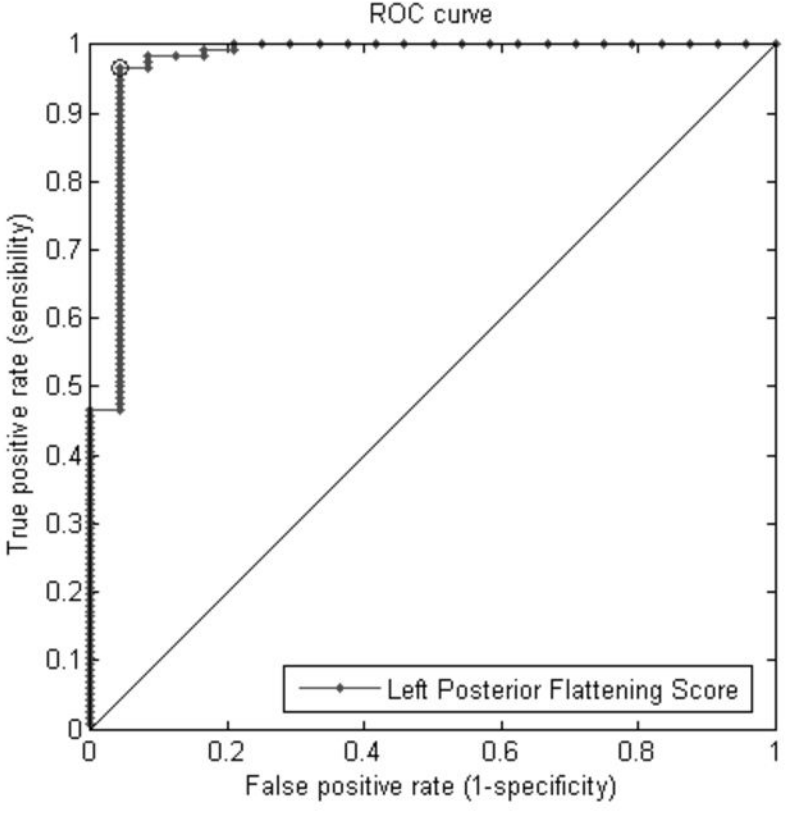

Left Posterior Flattening Score vs Expert Score
Expert Score
4
3
2
1
0
-1
-2
-3
-4
0
0.05
0.1
0.15
0.2
0.25
0.3
Left Posterior Flattening Score

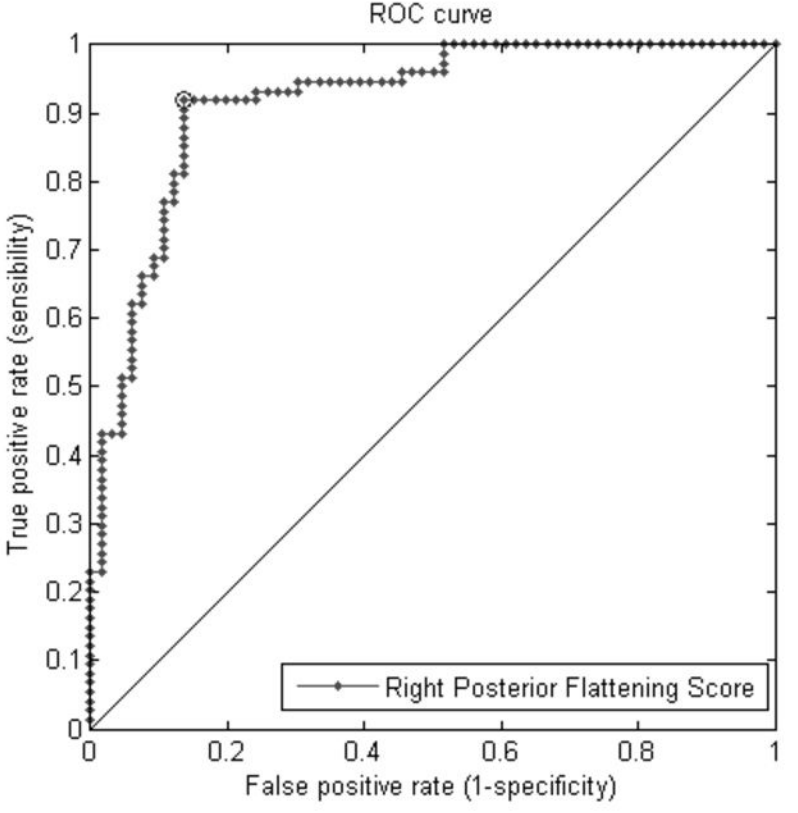

Right Posterior Flattening Score vs Expert Score
Expert Score
Right Posterior Flattening Score
4
3
2
1
0
-1
-2
-3
-4
0
0.05
0.1
0.15
0.2
0.25

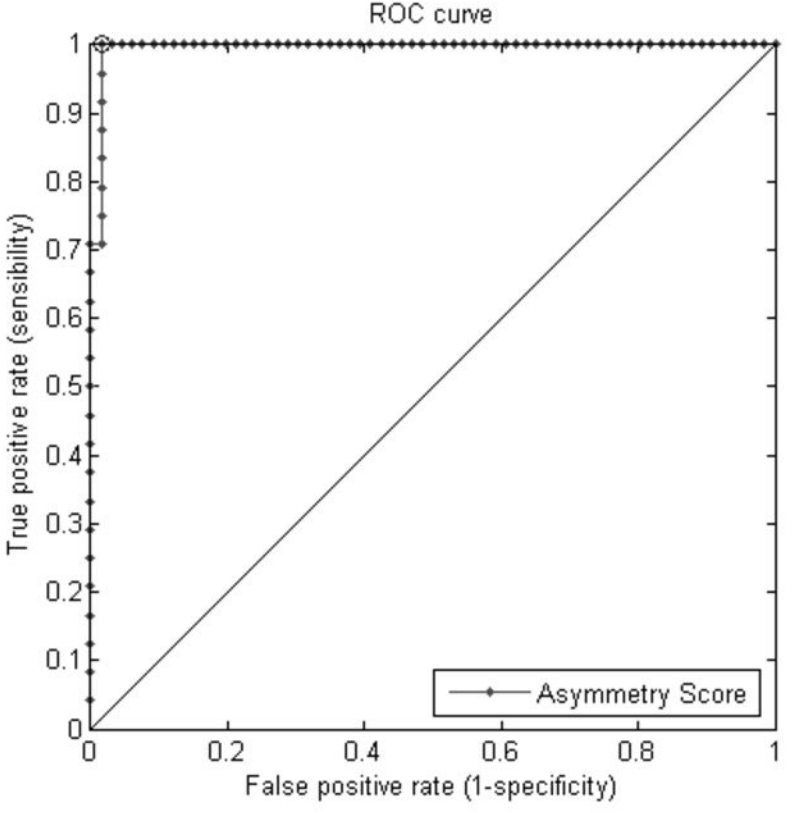

Asymmetry Score vs Expert Score
Expert Score
4
3
2
1
0
-2
-3
-4
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
Asymmetry Score

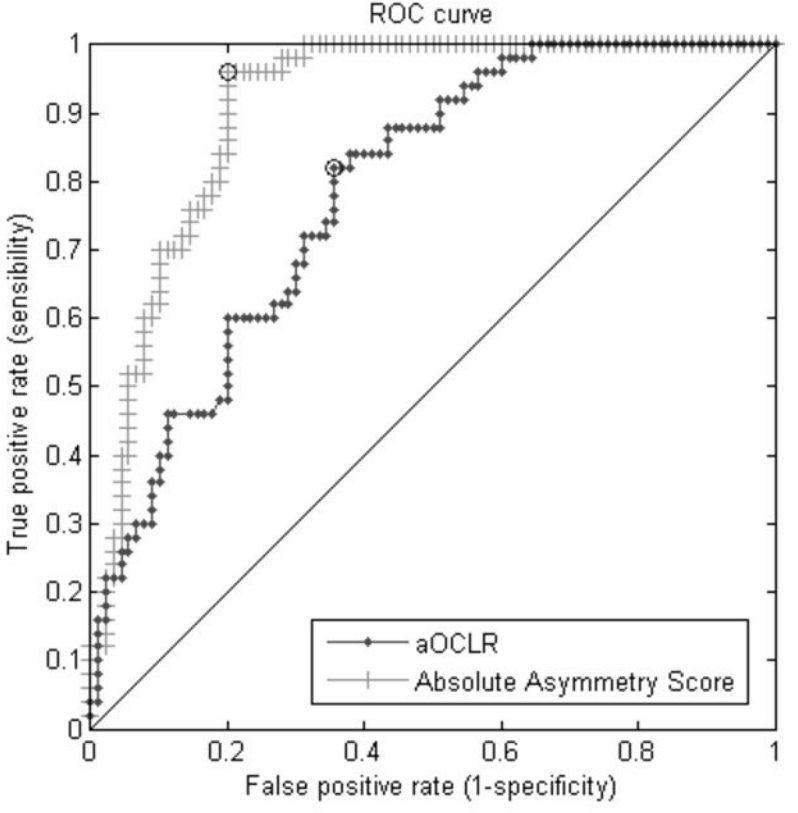

Absolute Asymmetry Score vs Expert Score
Expert Score
4
3.5
3
2.5
2
1.5
1
0.5
0
0
0.02
0.04
0.06
0.08
0.1
0.12
Absolute Asymmetry Score


Absolute Asymmetry Score (AAS) vs Oblique Cranial Length Ratio (OCLR)
AS threshold
OCLR threshold
OCLR
AAS
125
120
115
110
105
100
95
0
0.02
0.04
0.06
0.08
0.1
0.12
Expert Score 0
Expert Score 1
Expert Score 2
Expert Score 3


ROC curve
True positive rate (sensibility)
False positive rate (1-specificity)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1
Left Posterior Flattening Score


ROC curve
True positive rate (sensibility)
False positive rate (1-specificity)
Right Posterior Flattening Score
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1


ROC curve
True positive rate (sensibility)
False positive rate (1-specificity)
Asymmetry Score
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1


ROC curve
True positive rate (sensibility)
False positive rate (1-specificity)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1
aOCLR
Absolute Asymmetry Score