# CONCORDANCE OF THREE METHODS FOR PALPEBRAL FISSURE LENGTH MEASUREMENT IN THE ASSESSMENT OF FETAL ALCOHOL SPECTRUM DISORDER

*Meghan E Cranston[1,2], Aizeddin A Mhanni[1,2,3], Sandra L Marles[1,2,3], Albert E Chudley[1,2,3]*

*[1]Faculty of Medicine, [2]Department of Pediatrics and Child Health, [3]Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg, Manitoba, Canada*

## ABSTRACT

### Background
The assessment of individuals at risk of Fetal Alcohol Spectrum Disorders (FASD) includes the assessment of the craniofacial features that can result from prenatal alcohol exposure. The characteristic facial features of Fetal Alcohol Syndrome (FAS) consist of short palpebral fissures, smooth or flattened philtrum, and thin vermilion border of the upper lip. There are various methods for measuring palpebral fissure lengths (PFLs) and it can be challenging for clinicians to obtain reproducibly accurate measurements. The development of the FAS Facial Photographic Analysis Software by the University of Washington FAS Diagnostic and Prevention Network (DPN) is one such means of improving the accuracy and reproducibility in these measurements.

### Objectives
To assess concordance across three methods of PFL measurement: 1) a clear plastic handheld ruler, 2) blunt precision slide calipers, and 3) digital photometric photography (FAS Facial Photographic Analysis Software).

### Methods
The PFLs of 50 children (referred to the Clinic for Alcohol and Drug Exposed Children, CADEC) at Children's Hospital in Winnipeg and 50 adults from the University of Manitoba Medical Class of 2008 were measured once by a single clinician, using each of the three methods. The frequency and magnitude of discordance was tabulated. No method served as a gold-standard.

### Results
The PFLs ranged from 20 to 32 mm. The ruler and photometric measures were concordant in 42% of the subjects. When measures were discordant, half the ruler measures were larger and half were smaller. The caliper measure was concordant with the photometric and ruler measures on 18% and 24% of the subjects, respectively. When measures were discordant, the caliper measures were almost always larger than the photometric and ruler method (0.5 to 2.5 mm larger, 83% and 95% of the time, respectively). The presence of epicanthal folds did not appear to be a factor that contributed to discordance.

### Conclusion
This study demonstrates the challenge in measuring the PFL, even when a single trained clinician is involved. Factors that can contribute to error include the subject's willingness to cooperate, ability to tolerate placement of the tool close enough to the eye to obtain an accurate measure, and precision of the tool. When controlling for the clinician performing the measurements and the quality of the photographs, the ruler and photometric measures were most concordant. The caliper measures tended to measure larger than the ruler and photometric measures.

**Key Words**: *Fetal Alcohol Syndrome, Fetal Alcohol Spectrum Disorder, dysmorphology, palpebral fissure length, facial measurement tools, photographic analysis*

Fetal Alcohol Syndrome (FAS) is a permanent birth defect caused by maternal consumption of alcohol during pregnancy.[1] For over thirty years, FAS has been used to describe the constellation of abnormalities associated with ethanol teratogenesis. However, it has become apparent that the effects of prenatal alcohol exposure fall along a continuum from extreme to more subtle anomalies. Variable amounts of consumption and exposure at variable gestational timings offset by the multiple genetic and environmental influences produce a wide spectrum of phenotypic expression of the syndrome.[2] Fetal alcohol spectrum disorder (FASD) is a term that encompasses the wide range of physical, mental and behavioural effects that can occur when a person is exposed to ethanol during gestation. FASD has a total prevalence in the United States of 9.1 per 1000 births with its most severe form, fetal alcohol syndrome (FAS), and accounting for 1-3 per 1000 live births.[3]

A diagnostic evaluation for FASD focuses on four key diagnostic features: growth deficiency, a cluster of minor facial anomalies, central nervous system damage/dysfunction, and prenatal alcohol exposure.[4,5] The diagnostic features of the FAS facial phenotype include: 1) short palpebral fissures ($\leq 3^{rd}$ percentile or 2 SDs below the mean), smooth or flattened philtrum, and a thin vermilion border of the upper lip (rank 4 or 5 on the Lip-Philtrum Guide).[4,5] It is critical that the three facial features are measured accurately to document the magnitude of expression of the FAS facial phenotype. Palpebral fissure lengths (PFLs) remain problematic for clinicians to measure accurately and reproducibly.[5,6] Currently, no gold standard for PFL measurement exists. In clinical practice, PFL measurements are obtained using the clear plastic ruler and/or precision slide calipers. It is likely that disparities exist between measurements depending on, not only the training of the clinician and the cooperation of the individual being measured, but the measurement tool used. The development of the FAS Facial Photographic Analysis Software[7,8] has created an alternate method of measurement (pixel distance) for the palpebral fissure length using digital photographs and computerized image analysis tools.

The PFL is defined as the length between the endocanthion (inner corner of the eye) and exocanthion (outer corner of the eye). It has been traditionally measured with a clear plastic handheld ruler, and compared to published PFL growth chart norms by Hall et al.[9] for Caucasian individuals or Iosub et al.[10] for Black/Hispanic individuals. The accuracy of the PFL measure is influenced by the method used to measure it. The accuracy and reproducibility of the PFL measurement is critical because a difference of greater than one or more millimeters can impact the diagnostic classification of the FAS facial phenotype.

There is a need for standardized methods of measurement and training so every clinician will be able to reliably assess the FAS facial phenotype.

## METHODS

A total of 111 participants from two distinct groups were enrolled in the study. Eleven participants were excluded (10 children and one adult) due to poor photographs or lack of cooperation (see below).Thus 100 subjects completed the study. The first group included a prospectively collected cohort of 50 children seen through the Clinic for Alcohol and Drug Exposed Children (CADEC) at the Children's Hospital in Winnipeg, Manitoba who completed the study. The participants, aged 8 months to 13 years, were found to be at risk of FASD by their primary care physicians and referred to CADEC for a diagnostic assessment. Each child had both right and left PFLs measured as part of the routine FASD assessment at CADEC. The second group included a control group of 50 volunteer adults from the University of Manitoba's Faculty of Medicine Class of 2008 who completed the study. As a medical student requiring a certain high standard of academic achievement, each participant was assumed to be unaffected by prenatal exposure to alcohol. The exact age of the adult participant was not relevant, as it is accepted that PFLs do not change after 16 years of age.

Each participant had both right and left PFLs directly measured by one investigator (MEC) using a clear plastic handheld ruler and blunt slide (non-digital) calipers. PFLs were measured to the nearest 0.5mm with the instrument held as close as possible to the eye without touching the eye or eye lashes (Figure 1a,b).

**FIG. 1** Measuring palpebral fissure length. To measure PFL, the ruler (1a) or calipers (1b) are held in the horizontal plane closest to the eye. The exocanthion – endocanthion distance immediately below the eye is measured.

**FIG. 1 A**

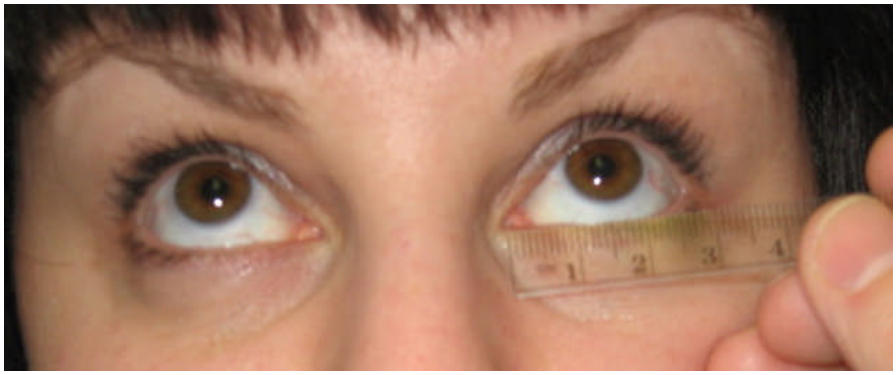
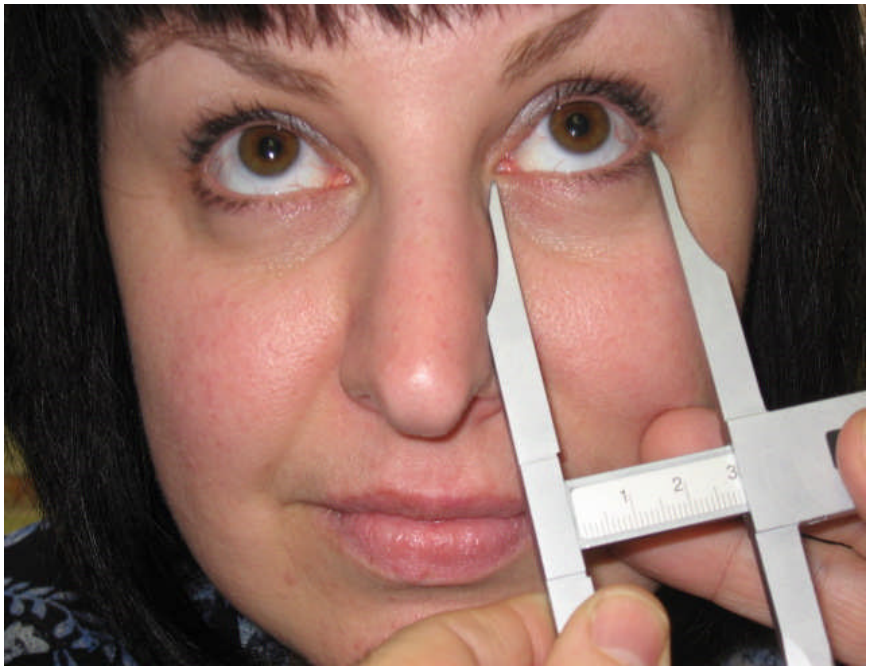


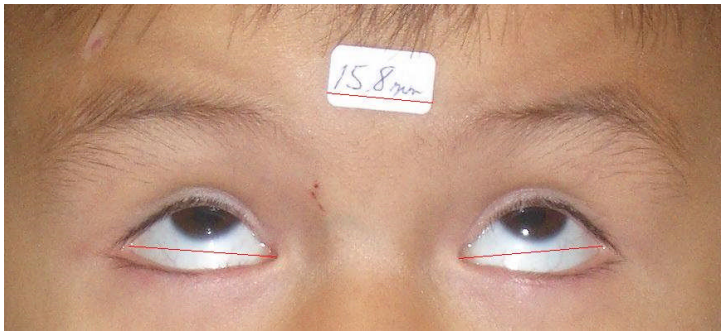
**FIG. 1 B**



The individual was asked to sit upright and open his/her eyes widely to allow accurate identification of the endocanthion and exocanthion landmarks. If the individual was observed to have epicanthal folds, the upper lid was gently pulled upwards just to expose the endocanthion of the eye before the measurement was taken.

Each participant had both PFLs measured by the digital photometric technique (FAS Facial Photographic Analysis Software developed by the University of Washington FAS DPN).[7] The investigator was trained in the use of the photometric technique by Astley who developed the method. A 15.8 mm (5/8") size adhesive was placed on the participant's forehead between the eyebrows to serve as an internal scale. The participant was asked to sit upright and to have a relaxed facial expression with the eyes fully open and mouth gently closed. A frontal view photographic image of the participant's face was obtained using a hand-held 4.0Mp digital camera. The right and left PFLs were measured by clicking the mouse pointer on the endocanthion and exocanthion landmarks and having the software compute the distance between the landmarks in pixel units (Figure 2).

**FIG. 2** Digital photometric photo image (FAS Facial Photographic Analysis Software). Lines plotted between the exocanthion and endocanthion allow software to measure and compute the PFL from the lines plotted on the internal scale (forehead adhesive).




The length of the internal measure of scale (paper adhesive) was also measured in units of pixels. The software is programmed to compute the actual size of the PFL to the nearest 0.1 mm based on the internal measure of scale. The program adjusts for the fact that the PFLs are off the midline and, thus, are foreshortened in the photo. If the endocanthion was not clearly visible in the photograph due to the presence of epicanthal folds, we estimated the anatomical location of the endocanthion by extending the arc of the upper and lower palpebral fissure to the point of intersection.

To best achieve intra-rater reliability, each participant had their PFLs measured first by the ruler or calipers in a random order. The clinician was not blinded to results of either measurement. The photos were taken at the same visit as the clinical measurements; however, the photometric analysis was performed two to three weeks following the visit making it extremely difficult for the clinician to reasonably remember previous measures for each participant.

Participants who were uncooperative or who did not tolerate the tool close enough to the eye for an accurate measurement were excluded from the study. Also, participants whose photographs appeared blurry or rotated more than 8 degrees to the left or right were not included in the study. A total of 11 participants (10 children and 1 adult) were excluded from the prospectively collected cohort.

The mean of the right and left PFLs were computed for each participant for all three methods of measurement. The prevalence of concordance and the magnitude of discordance between the three measures for each participant were documented. An analysis of variance (ANOVA) test was employed select for the presence of epicanthal folds and determine overall

significance in altering PFL measurements. This study was approved by the University of Manitoba Bannatyne Campus Health Research Ethics Board.

## RESULTS

The palpebral fissure lengths ranged from 20 to 32 mm for 100 child and adult participants. The ruler and photometric measures were concordant in 42% of the subjects (Tables 1 and 2). When measures were discordant, 50% of the ruler measures were larger and 50% were smaller. The caliper measure was concordant with the photometric and ruler measures on 18% and 24% of the subjects, respectively. When measures were discordant, the caliper measures were almost always larger than the photometric and ruler method (0.5 to 2.5 mm larger, 84% and 95% of the time, respectively). An accuracy of 1 or more mm could impact diagnostic classification of the FAS facial phenotype. This magnitude of difference was observed in 27% to 45% of the paired measures.

**TABLE 1** Prevalence of concordant and magnitude of discordant measures between measures obtained from the caliper, ruler, photometric methods for each participant (n=100).

**Table 1**

| Difference in PFL (mm) | Caliper minus Ruler | Caliper minus Photometric | Ruler Minus Photometric |
|---|---|---|---|
| | N = 100 | N = 100 | N = 100 |
| -2.5 mm | 0 | 0 | 0 |
| -2 mm | 0 | 0 | 0 |
| -1.5 mm | 0 | 1 | 3 |
| -1 mm | 1 | 3 | 10 |
| -0.5 mm | 3 | 9 | 16 |
| Exact match 0 mm | 24 | 18 | 42 |
| +0.5 mm | 37 | 28 | 15 |
| +1 mm | 31 | 24 | 10 |
| +1.5 mm | 3 | 11 | 2 |
| +2 mm | 1 | 5 | 2 |
| +2.5 mm | 0 | 1 | 0 |

**TABLE 2** Comparison of discordant measures between measures obtained from the caliper, ruler, and photometric methods for each participant (n=100).

**Table 2**

| | Caliper - Ruler | Caliper - Photometric | Ruler - Photometric |
|---|---|---|---|
| Total No. of Measures N | 100 | 100 | 100 |
| Concordant Measures N<br>% | 24<br>24% | 18<br>18% | 42<br>42% |
| Discordant Measures N<br>% | 76<br>76%<br>(36%≥ 1mm) | 82<br>82%<br>(45% ≥ 1mm) | 58<br>58%<br>(27%≥ 1mm) |
| Larger N<br>% | 72<br>95% | 69<br>84% | 29<br>50% |
| Smaller N<br>% | 4<br>5.2% | 13<br>17% | 29<br>50% |

Of the total 100 participants, 32 were observed to have epicanthal folds. When selecting for the presence of epicanthal folds in our analysis, the anomaly was not found to have overall significance in altering the PFL measurements (p=0.0709).

The PFL measurements obtained from the children at risk for FASD follow the trajectory of the standard PFL growth chart developed and published by Hall et al.[9] However, our measurements consistently fall between 0.5 and 1 standard deviation (SD) below the published means. The mean PFL published for Caucasian adults is 30.7 mm (SD 1.4) compared to our adult Caucasian PFL findings of 29.1 mm (SD 1.32), 29.3 mm (SD 1.33), and 30.0 mm (SD 1.26), when obtained by the photometric, ruler, or caliper methods, respectively. Consistent with our findings for the children, the adult control group demonstrated a mean PFL measurement 0.5 to 1 SD below the published adult mean by Hall et al.[9]

## DISCUSSION

Maternal consumption of ethanol is the leading known cause of mental deficiency in North America and produces an identifiable pattern of altered growth and morphogenesis.[6] It is necessary that we use most accurate methods to measure the magnitude of FAS appearance in the assessment process. In the spectrum of this altered growth and morphogenesis, short PFL is the most consistent physical feature.[11] From the results of 100 individuals, it is quite clear that inconsistencies exist in the PFL measurements obtained depending on the method of measurement used. The ruler and photometric technique yielded a higher concordance in measurements, while the calipers consistently yielded larger measurements. It is unclear why the calipers measure larger than the other methods. The digital photometric technique was developed as a precise screening tool and diagnostic aid to minimize the possibility of human error in PFL measurements using the ruler or calipers.[8] The photometric analysis produces similar PFL measurements to the ruler whereas the calipers measure larger in this study. The plastic ruler has been the standard clinical tool in most clinics because it is inexpensive and readily available. The use of the ruler is highly dependent on the training of the clinician. Even though PFL measurement seems quite straight forward, very few professionals are trained in the assessment of the face. The use of the ruler for this purpose requires training and experience. It is paramount that clinicians confirm the accuracy of their PFL measurements. In the future, the use of the photometric analyses for the screening and diagnosis of FASD may become more broadly used. The photographic analysis has the ability to transfer and store data for centralized analysis maximizing consistency of interpretation and may present new opportunities for screening and telehealth diagnosis, making it possible for many more children at risk of FASD who may live a distance from diagnostic centres to receive timely diagnoses.

We compared the mean PFLs obtained by our measurements to the standard PFL growth chart developed and published by Hall et al.[9] The PFL measurements obtained for the children at risk for FASD did follow the trajectory of the chart; however, they consistently fell 0.5 to 1 SD below the published mean. This might be expected, as this group of children was likely exposed to alcohol during gestation. However, it is questionable whether this difference is solely due to the fact that the cohort are alcohol exposed and as a result have short PFLs, or whether Hall's chart overestimates what is actually seen in the normal population. The latter argument is supported by the fact that the healthy adult control group also demonstrated a mean PFL measurement 0.5 to 1 SD below the mean. The data suggest that the currently published growth charts overestimate the PFLs of children and adults.

There are a number of standardized PFL growth norm charts for Caucasians Those published in Hall et al.[9], combined data, including those from Farkas[12] and Thomas et al.[13] The measurement norms in the chart by Thomas et al. are almost identical with those by Hall et al. until the age of 6 years; however, they are substantially smaller when compared to norms by Farkas after the age of 6 years.

The chart by Thomas et al. involved a mathematical model predicting growth trajectory of PFLs for children over the age of 6 years and is not based on actual measurements of children. It appears that norms by Hall et al. were gathered from Farkas' data for children over the age of 6

years. Farkas' data were obtained with slide calipers, whereas Hall's data for children under the age of 6 years were measured by handheld rulers or calipers depending on the source of the data. This may explain the discrepancy in the charts and the reason for higher estimates for PFL in older children when compared to clinical measurements using the handheld ruler. Norms by Hall et al. and Farkas for 16 year old Caucasian PFLs yield a mean of 30.7 mm, whereas Thomas et al. reports a mean of 27 mm (a difference of 3 SD). Since the published 'standardized norms' vary by as much as 1-3 SD, it is clear that more representative population norms are required if clinicians are to reliably and accurately diagnose FASD individuals, especially for those over the age of 6 years.

One recently published study compared the photographic method with direct ruler measurements for PFLs in 40 children.[14] The authors found that the photographic method showed lower measurements for PFLs than direct ruler measurements, but this was only statistically different for children under the age of 4 years. The authors did not compare with the caliper method. However, for the older children, the results are in agreement with our study results.

A key limitation of this study is the absence of a gold standard for PFL measurement. With the absence of a gold standard measurement of PFL for each participant, our study cannot comment on the accuracy of any of the methods of PFL measurement that we compared. Our study was potentially biased by the fact that the first direct PFL measure (ruler/calipers) could not be masked to the clinician before the second measure was obtained the same day. We may have further minimized this bias by measuring each participant with one tool on two separate visits. The goal would be to have a reasonable length of time between visits to prevent clinician recall of the previous measure. Although we observed no significant difference in measurements obtained from participants with and without epicanthal folds, no measurement can, theoretically, be accurately reproduced when the soft tissue of the palpebral fissure is manually distorted prior to measurement. Clinicians must agree on how to estimate the location of the endocanthion when the landmark is obscured by epicanthal folds. It is clear that PFL measures vary depending on the method of measurement used. Our study demonstrates the challenge in measuring the PFL, even when a single trained clinician is involved. There are a multitude of factors that can contribute to error including the participant's willingness to cooperate, ability to tolerate placement of the tool close enough to the eye to obtain an accurate measure, and precision of the tool. When controlling for the clinician performing the measurements (level of training and hand-eye coordination) and the quality of the photographs, the ruler and photometric measures were most concordant. When the caliper measures were discordant from the ruler and photometric methods, the calipers tended to measure larger that the other methods. However, our study was not designed to compare the accuracy of the various methods of PFL measurement due to the fact that a gold standard for PFL measurement has not been defined.

Our study demonstrates that PFL variability does not only exist between different measurement tools and clinicians, but also among the age-matched reference norms to which measured PFLs are compared. Since the published standardized norms vary, it is apparent that we need more reliable normative data in order to confidently achieve an FAS diagnosis.

**Corresponding Author:** *achudley@hsc.mb.ca*

***Acknowledgements***

We thank the children and families who agreed to participate in this study. We thank Dr. Susan Astley, Department of Epidemiology, University of Washington for providing extensive insight, review, and feedback. We also thank Dr. M. Cheang from the Biostatistical Consulting Unit at the University of Manitoba for her work. This research was facilitated by members of the Clinic for Alcohol and Drug Exposed Children (CADEC) at the Children's Hospital of Winnipeg. This research was supported by a grant from the Manitoba Institute of Child Health. This research represents partial fulfillment of the B.Sc. (Medicine) degree, Faculty of Medicine, University of Manitoba, awarded to MEC.

***Conflict of Interest*:** None declared

**REFERENCES**


1. Jones KL, Smith DW, Ulleland CN, Streissgnuth AP. Pattern of malformation in offspring of chronic alcoholic mothers. Lancet 1973;1(7815):1267-71.
2. Clarren SK, Smith DW. The fetal alcohol syndrome. New Eng J Med 1978;298(19):1063-7.
3. Sampson PD, Streissguth AP, Bookstein FL, Little RE, Clarren SK, Dehaene P, et al. Incidence of fetal alcohol syndrome and prevalence of alcohol-related neurodevelopmental disorder. Teratology 1997;56(5):317-26.
4. Chudley AE, Conry J, Cook JL, Loock C, Rosales T, LeBlanc N, and the Public Health Agency of Canada's National Advisory Committee on Fetal Alcohol Spectrum Disorder. Fetal alcohol spectrum disorder: guidelines for diagnosis. CMAJ 2005;172(5suppl):S1-S21.
5. Astley SJ. Diagnostic Guide for Fetal Alcohol Spectrum Disorders: the 4-Digit Diagnostic Code, 3rd edn. Seattle, WA: University of Washington Publication Services, 2004.
6. Astley SJ, Clarren SK. Measuring the facial phenotype of individuals with prenatal alcohol exposure: correlations with brain dysfunction. Alcohol Alcohol 2001;36(2):147-59.
7. Astley SJ. Fetal Alcohol Syndrome Facial Photographic Analysis Software, Version 1.0.0, University of Washington, Seattle WA, 2003.
8. Astley SJ, Clarren SK. A case definition and photographic screening tool for the facial phenotype of fetal alcohol syndrome. J Pediatr 1996;129(1):33-41.
9. Hall JG, Froster-Iskenius UG, Allanson JE, editors. Handbook of normal physical measurements. Oxford: Oxford University Press; 1989. pp. 149-50.
10. Iosub S, Fuchs M, Bingol N, Stone R, Gromisch D, Wasserman E. Palpebral fissure length in black and hispanic children: Correlation with head circumference. Pediatr 1985;75(2):318-20.
11. Jones KL, Hanson JW, Smith DW. Palpebral fissure size in newborn infants. J Pediatr 1978;92(5):787.
12. Farkas LG. Anthropometry of the head and face. 2nd ed. New York: Raven Press; 1994.
13. Thomas IT, Gaitantzis YA, Frias JL. Palpebral fissure length from 29 weeks gestation to 14 years. J Pediatr 1987;111:267-8.
14. Avner M, Henning P, Koren G, Nulman I. Validation of the facial photographic method in fetal alcohol spectrum disorder screening and diagnosis. JFAS Int 2008; 4:e20.