A case definition and photographic screening tool for the facial phenotype of fetal alcohol syndrome

Susan J. Astley, PhD, and Sterling K. Clarren, MD

From the Department of Epidemiology, School of Public Health and Community Medicine, and the Department of Pediatrics, School of Medicine, University of Washington, Seattle

Objectives: **The purpose of this study was to demonstrate that a quantitative, multivariate case definition of the fetal alcohol syndrome (FAS) facial phenotype could be derived from photographs of individuals with FAS and to demonstrate how this case definition and photographic approach could be used to develop efficient, accurate, and precise screening tools, diagnostic aids, and possibly surveillance tools.**

Study design: **Frontal facial photographs of 42 subjects (from birth to 27 years of age) with FAS were matched to 84 subjects without FAS. The study population was randomly divided in half. Group 1 was used to identify the facial features that best differentiated individuals with and without FAS. Group 2 was used for cross validation.**

Results: **In group 1, stepwise discriminant analysis identified three facial features (reduced palpebral fissure length/inner canthal distance ratio, smooth philtrum, and thin upper lip) as the cluster of features that differentiated individuals with and without FAS in groups 1 and 2 with 100% accuracy. Sensitivity and specificity were unaffected by race, gender, and age.**

Conclusions: **The phenotypic case definition derived from photographs accurately distinguished between individuals with and without FAS, demonstrating the potential of this approach for developing screening, diagnostic, and surveillance tools. Further evaluation of the validity and generalizability of this method will be needed. (J Pediatr 1996;129:33-41)**

Fetal alcohol syndrome is a permanent birth defect syndrome caused by maternal consumption of alcohol during pregnancy. FAS is characterized by cognitive and behavioral dysfunction, a unique cluster of minor facial anomalies, and prenatal or postnatal growth deficiency.[1] FAS is the leading known cause of mental retardation in the Western World,[2] with an estimated incidence of 1 to 3 per 1000 live births.[3] Individuals with FAS endure lifelong physical, intellectual, cognitive, and behavioral disabilities. These disabilities are often compounded by secondary emotional and behavioral disabilities such as low self-esteem, depression, school failure, and criminality when the syndrome fails to be diagnosed. These secondary disabilities come at a high cost to the individual, his or her family, and society and can be reduced by early diagnosis and receipt of appropriate intervention.[4, 5]

See commentary, p. 3.

D-score	Discriminant score
FAS	Fetal alcohol syndrome

Efforts to prevent FAS and its associated secondary disabilities are currently stymied by the lack of efficient and effective surveillance, screening, and diagnostic tools. Devel-

Supported in part by the Chevez FAS Fund at Children's Hospital and Medical Center, Seattle, Wash.

Submitted for publication Jan 16, 1996; accepted April 5, 1996.

Reprint requests: Susan J. Astley, PhD, Children's Hospital and Medical Center, 4800 Sand Point Way N.E., Seattle WA 98105.

0022-3476/96/$5.00 + 0 **9/20/74002**

opment of these tools have, in turn, been stymied by the lack of an effective FAS case definition.[6] An ideal case definition for screening and surveillance of FAS would focus on the minimum cluster of features unique to FAS that are amenable to accurate, precise, and efficient measurement. The FAS facial phenotype is characterized by a cluster of minor facial anomalies that include small palpebral fissures, a smooth philtrum, and a thin upper lip.[1] Criteria have never been established regarding how small, how smooth, or how thin these features must be, nor have criteria been established as to how many of these features must be present.

The goals of this study were to demonstrate (1) that an objective, quantitative, multivariate case definition of the FAS facial phenotype could be derived from facial photographs and (2) that this methodologic approach could be used to develop highly efficient, accurate, and precise screening tools, surveillance tools, and diagnostic aids for FAS.

METHODS

Overview. Frontal facial photographs of 42 subjects with FAS (birth to 27 years of age) were pair matched on age, race, and gender to the frontal facial photographs of two randomly selected subjects without FAS (n = 84). The 126 patients were randomly divided into two groups (n = 63 per group) balanced on age, gender, and race. Stepwise discriminant analysis was used to identify the facial feature(s) that best differentiated the subjects with and without FAS in group 1. The multivariate discriminant equation generated from the discriminant analysis in group 1 served as both the FAS phenotypic case-definition and the method (or screening tool) by which risk of FAS was assessed among patients in group 2. Application of the screening tool to group 2 served as an opportunity to test the tool's validity.

Study population and photographic quality. Photographs of patients with FAS and of control subjects were selected from among 1110 frontal facial images of 740 clinical and research subjects; the images had been stored in a computerized image database. This study was conducted with the approval of the University of Washington and Children's Hospital and Medical Center Human Subjects Divisions.

Computerized images of subjects with FAS have been collected over the years from syndrome diagnosis textbooks, the medical literature, colleagues, and patients examined by one of the authors at the Centers for Disease Control and Prevention–sponsored University of Washington FAS Clinic. Each patient with FAS who was selected for this study was examined by a clinician with recognized expertise in the diagnosis of FAS and was believed, by the clinician, to have the facial phenotype of FAS at the time the photograph was taken. There were no age, race, or gender restrictions placed on selection of FAS cases.

Each patient with FAS was matched, on gender, race, and age when the photograph was taken (within 2 years), to two control subjects confirmed not to have a diagnosis of FAS. Control subjects may or may not have had documented prenatal alcohol exposure. The study control subjects were selected from a pool of 560 subjects that included patients seen at the University of Washington FAS Clinic and subjects who had participated in previous morphometric studies. Each pair of control subjects was randomly selected from among all control subjects meeting the matching criteria for each FAS case.

Photographs had to meet the following criteria for inclusion into the study: (1) the camera was aligned in the Frankfort horizontal plane with minimal left-to-right rotation (i.e., both planes of rotation were within ±5 degrees by visual inspection), (2) the subject had a relaxed facial expression with eyes fully open and lips gently closed, and (3) the image had adequate exposure and focus to allow accurate measurement of facial features.

Computer images and facial measures. All photographs were captured at 640 × 480 pixel resolution on a 256-unit gray scale with OPTIMAS (Optimas Corp., Edmonds, Wash.) image acquisition and enhancement software. The images were saved in tag-image-file (TIF) format. To reduce measurement bias, we masked the FAS case-control status of each photograph by cropping two separate images from each original TIF image: (1) a computer image that included just the eyes and (2) a computer image that included just the philtrum and mouth. All facial measurements were collected from these cropped images.

The facial measures recorded from each subject are listed and defined in Table I. Emphasis was placed on measuring features previously identified as being specific to FAS[7] and features that could be reliably measured from photographs. Right and left palpebral fissure lengths and inner canthal distance were measured with the distance measurement tool in the OPTIMAS software. The photographs did not contain internal measures of scale; therefore a reduced palpebral fissure length/inner canthal distance ratio was used as a proxy measure for a short palpebral fissure. Direct measures obtained from patients with a diagnosis of FAS at the University of Washington FAS Clinic have confirmed the presence of short palpebral fissures relative to normal inner canthal distances,[7] supporting the validity of the proxy measure. The phenotypic expressions of philtrum smoothness and upper lip thinness were recorded on 5-point Likert scales (Fig. 1). Upper lip thinness was also recorded on a continuous scale with an objective measure of shape called circularity (or perimeter^2/area). The circularity of a circle is 4*Pi (about 12.57), which is theoretically the smallest value that this measure can have. As an object tends toward the shape of a line (or thinness), the circularity tends toward infinity. Philtrum contour was also measured objectively by recording pixel luminosity on a 256 continuous gray scale. With

the OPTIMAS software, a line was drawn horizontally across the philtrum, centered between the upper lip and subnasion, with a length equal to the distance from the left to the right corners of the mouth and a width 20% of the vertical distance between the upper lip and the subnasion. The length of the line was divided into 100 units of equal size. Pixel luminosity was averaged over each of the 100 units and plotted to portray the gray scale variation (or shadows and highlights cast by the philtrum's ridges) across the philtrum (Fig. 2). The darkest luminosity in the philtrum furrow was subtracted from the brightest luminosity at a philtrum ridge to generate a measure of philtrum smoothness. The more deeply furrowed the philtrum, the greater the contrast in the luminosity of the ridge relative to the furrow. Upper-lip thinness and philtrum smoothness were recorded on both Likert ordinal scales and objective continuous scales to assess the reliability and utility of each. The Likert scale has the advantage of being technologically simple, meaning that the feature can be scored by hand if a computer is unavailable. The disadvantage is a lower degree of accuracy and precision. In direct contrast, the continuous measures of circularity and luminosity have the advantage of total objectivity, resulting in a high degree of accuracy and precision, but require access to appropriate software for derivation.

Discriminant analysis. Stepwise discriminant analysis (maximizing the Wilk lambda) was used to identify the facial feature(s) that best differentiated patients with and without FAS in group 1. Prior probability of FAS was set equal to the prevalence in the study sample (33%). The unstandardized canonical discriminant function coefficients were computed to derive the discriminant equation for calculation of each subject's discriminant score. The D-score was used to *predict* each subject's diagnostic classification (FAS, not FAS). The D-score distribution for the subjects with and without FAS in group 1 were plotted to identify the D-score cutoff value that resulted in the most accurate diagnostic prediction (i.e., had the highest level of sensitivity and specificity). Sensitivity is the proportion of patients with FAS who are correctly screened as having FAS. Specificity is the proportion of patients without FAS who are correctly screened as *not* having FAS. Sensitivity and specificity were computed by comparing each subject's true clinical diagnosis with their predicted diagnosis derived from the discriminant equation.

For cross validation the discriminant equation and D-score cutoff value derived from group 1 were applied to group 2. Sensitivity and specificity for group 2 were computed by comparing each subject's true diagnosis with their predicted diagnosis.

RESULTS

The total population of 42 subjects with FAS and 84 control subjects were successfully balanced on gender, race, and age at the time the photograph was taken (Table II). The age distribution of the study population was as follows: birth to 2 months (n = 1), 3 to 12 months (n = 4), 1 to 5 years (n = 43), 6 to 10 years (n = 46), 11 to 15 years (n = 22), 16 to 20 years (n = 6), and 20+ years (n = 4). No two subjects with FAS had an identical pattern of facial features. The variation of phenotypic expression across the 42 subjects with FAS relative to the 84 control subjects is displayed in Fig. 3.

Table I. Facial measures recorded from the computerized frontal facial photographs

Facial feature	Description (units of measurement)
Eye region	
Palpebral fissure lengths	Distance between outer and inner canthi of right and left eyes (centimeters of computer monitor screen)
Inner canthal distance	Distance between right and left inner canthi (centimeters of computer monitor screen)
Mouth region	
Philtrum smoothness	Area between upper lip and subnasion, with focus on presence of midline vertical furrow bordered by two vertical ridges (Fig. 1)
	5-Point Likert ordinal scale (1 deeply furrowed, 2 somewhat furrowed, 3 mid range, 4 somewhat smooth, 5 very smooth)
	Pixel luminosity: contrast between philtrum's ridges and furrow—the lower the contrast, the smoother the philtrum (0 to 255 continuous gray scale: 0 = white, 255 = black)
Upper lip thinness	Upper lip demarcated by its vermilion border (Fig. 1)
	5-Point Likert ordinal scale (1 very thick, 2 somewhat thick, 3 mid-range, 4 somewhat thin, 5 very thin)
	Circularity—the larger the circularity, the thinner the upper lip (perimeter^2/area)

Discriminating facial features and the influence of measurement scale. Stepwise discriminant analysis selected all three facial features (palpebral fissure length/inner canthal distance ratio, philtrum smoothness measured on a Likert scale, and upper lip thinness measured on the continuous scale of circularity) as the cluster of features that best differentiated the patients with and without FAS. When the ordinal and continuous scales used to measure philtrum smoothness and upper lip thinness were interchanged in the discriminant equation, sensitivity and specificity were min-

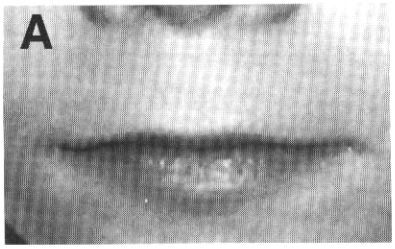
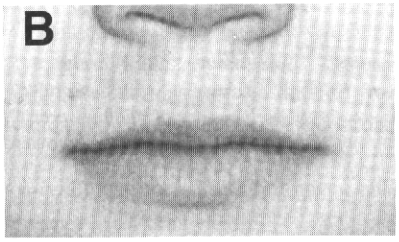
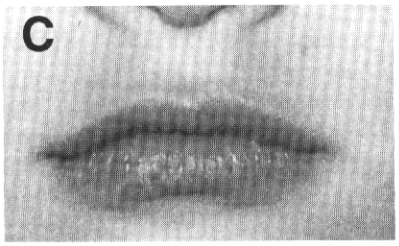
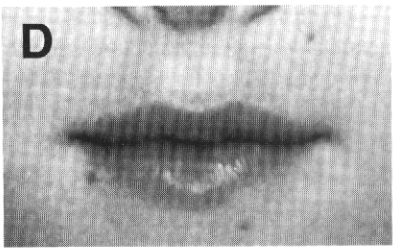
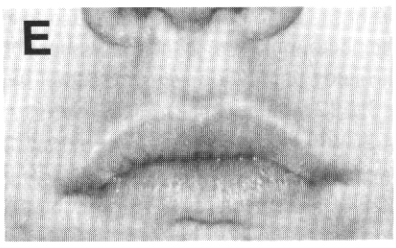
Philtrum and upper lip	Philtrum Likert score	Upper lip Likert score	Philtrum luminosity	Upper lip circularity
 A	5	5	0	178.6
 B	4	4	2	72.0
 C	3	3	4	55.1
 D	2	2	7	50.0
 E	1	1	15	44.5

Fig. 1. Pictorial examples of the 5-point Likert ordinal scales used to rank upper lip thinness and philtrum smoothness. The corresponding continuous measures of upper lip circularity (perimeter^2/area of vermilion border) and philtrum luminosity (contrast in pixel luminosity between the philtrum's ridge and furrow, with luminosity measured on a 256-unit gray scale) are also presented.

imally influenced. The discriminant classification and cross validation results are presented below for each measurement scale.

Philtrum and upper lip measured on a Likert scale. When palpebral fissure length/inner canthal distance ratio, philtrum smoothness measured on a Likert scale, and upper lip thinness measured on a Likert scale were entered into the discriminant equation derived from group 1, a D-score greater than 0.7 differentiated the subjects with and without FAS with 100% sensitivity and specificity. When this dis-

Table II. Demographic profile of subjects with and without FAS

Characteristic	Control group (n = 84)	FAS group (n = 42)
Age (yr)		
Mean (SD)	7.9 (4.7)	6.5 (5.2)
Minimum-maximum	1.5-26.7	0.0-24.2
Gender		
Girls: n(%)	28 (33)	14 (33)
Race: n(%)		
White	64 (77)	32 (77)
Black	4 (5)	2 (5)
American Indian	10 (12)	5 (12)
Alaskan Native	2 (2)	1 (2)
Asian	2 (2)	1 (2)
Hispanic	2 (2)	1 (2)
Other syndromes: n(%)		
Williams syndrome	1 (1)	0 (0)
Dubowitz syndrome	1 (1)	0 (0)
Aarskog syndrome	1 (1)	0 (0)
Marfan syndrome	1 (1)	0 (0)

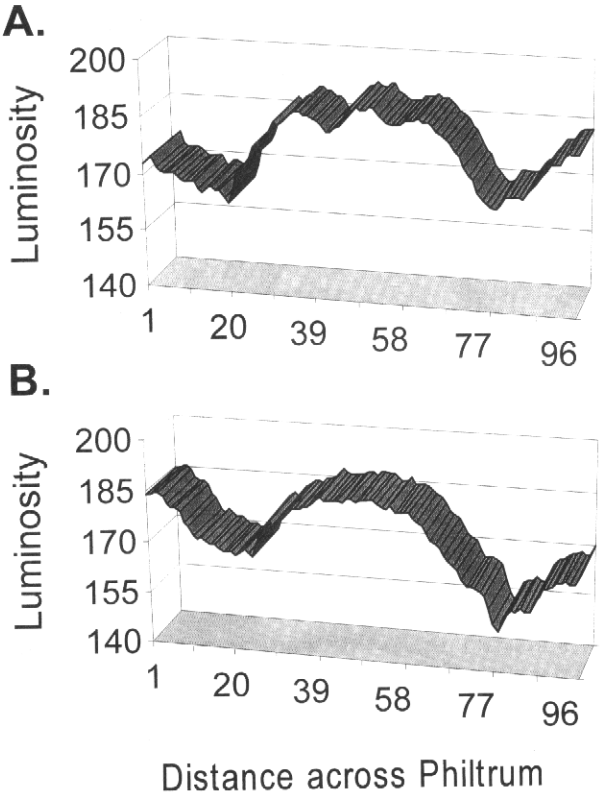


Fig. 2. Plot of pixel luminosity on a 256-unit gray scale along a line drawn horizontally across the philtrum from the left corner of the mouth (x-axis = 1) to the right corner of the mouth (x-axis = 100). A luminosity of 0 = black, and a luminosity of 255 = white. **A,** Example of a deeply furrowed philtrum (Likert scale = 1). **B,** Example of a very smooth philtrum (Likert scale = 5).

criminant equation and D-score cutoff value were applied to group 2 for cross validation, subjects from group 2 were also differentiated with 100% sensitivity and specificity.

The discriminant equation (equation 1), derived across all 126 subjects, was as follows: D-score = 0.7408 – 5.7337 (palpebral fissure length/inner canthal distance ratio) + 1.1677 (philtrum Likert score) + 0.1587 (upper lip Likert score). A D-score greater than 0.8 was the cutoff value for classifying a subject as having FAS on the basis of screening (sensitivity = 100%, specificity = 100%, and overall accuracy = 100%). Discriminant equation 1 explained 100% of the total variance (chi-square value [3, 126] = 224; p = 0.0000).

Philtrum and upper lip measured on continuous scales. When palpebral fissure length/inner canthal distance ratio, philtrum smoothness measured on the continuous scale of luminosity, and upper lip thinness measured on the continuous scale of circularity were entered into the discriminant equation derived from group 1, a D-score less than –0.5 differentiated the subjects with and without FAS, with 95% sensitivity and 100% specificity. When this discriminant equation and D-score cutoff value were applied to group 2 for cross validation, subjects from group 2 were differentiated with 100% sensitivity and 93% specificity.

The discriminant equation (equation 2) derived across all 126 subjects was as follows: D-score = –6.4719 + 8.6104 (palpebral fissure length/inner canthal distance ratio) + 0.0767 (philtrum luminosity) – 0.0145 (upper lip circularity). A D-score less than –0.5 = the cutoff value for classifying a subject as having FAS on the basis of screening (sensitivity = 99%, specificity = 95%, and overall accuracy = 98%).

Discriminant equation 2 explained 100% of the total variance (chi-square value [3, 126] = 130; p = 0.0000).

Philtrum measured on a Likert scale, upper lip measured on a continuous scale. When palpebral fissure length/inner canthal distance ratio, philtrum smoothness measured on a Likert scale, and upper lip thinness measured on a continuous scale of circularity were entered into the discriminant equation derived from group 1, a D-score greater than 0.7 differentiated the subjects with and without FAS with 100% sensitivity and specificity. When this discriminant equation and D-score cutoff value were applied to group 2 for cross validation, subjects from group 2 were also differentiated with 100% sensitivity and specificity.

The discriminant equation (equation 3), derived across all 126 subjects was as follows: D-score = 1.1075 – 6.0082 (palpebral fissure length/inner canthal distance ratio) + 1.1448 (philtrum Likert score) + 0.0066 (upper lip circularity). A D-score greater than 0.7 = the cutoff value for

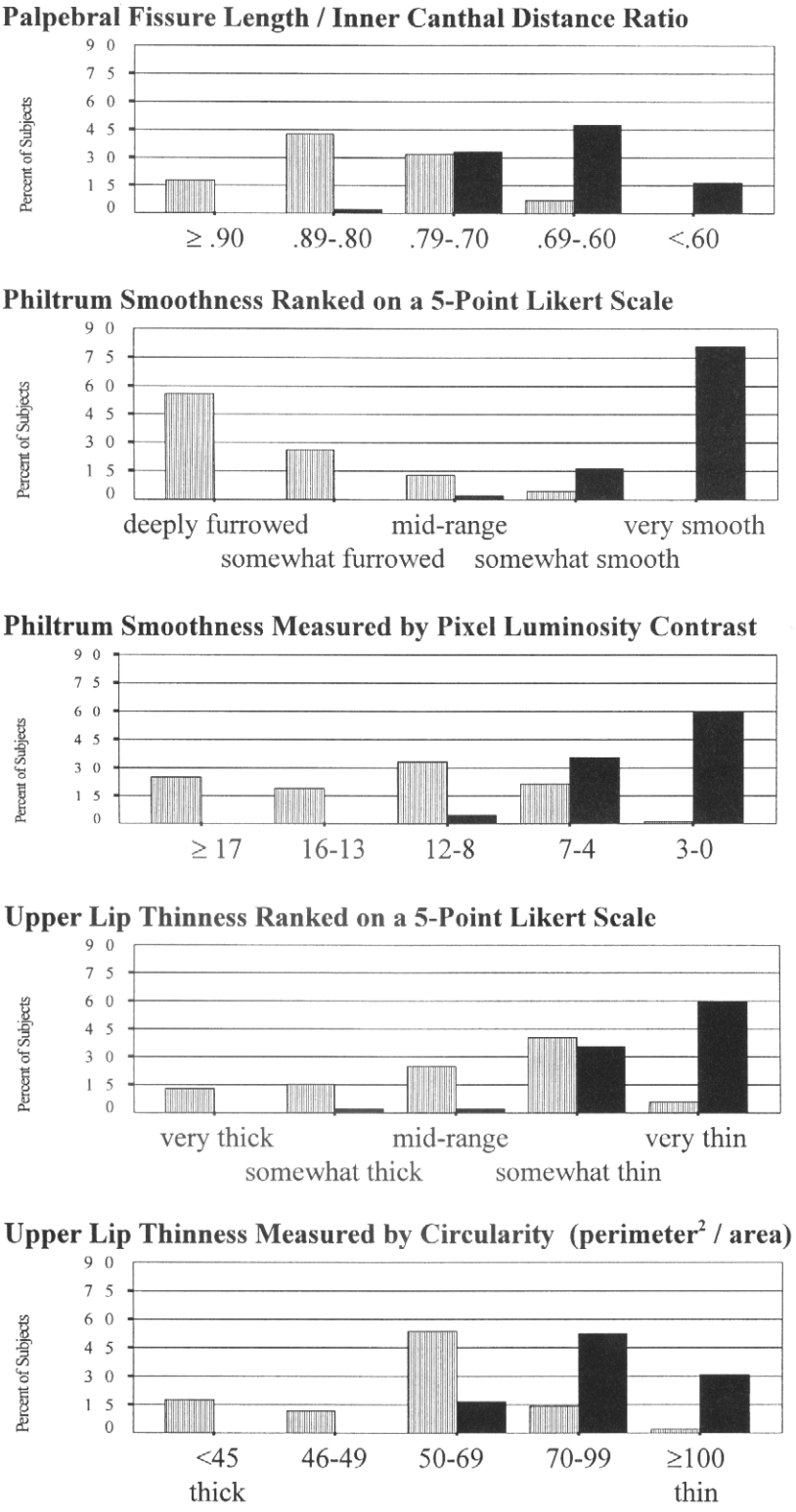


Fig. 3. Variation in expression of facial features among the 42 subjects with *(black bars)* FAS and 84 subjects without *(gray bars)* FAS. The circularity measure of upper lip thinness reflects the perimeter^2/area of the vermilion border. Philtrum luminosity contrast reflects the contrast in pixel luminosity between the philtrum's ridge and furrow, with luminosity recorded on a 256-unit gray scale.

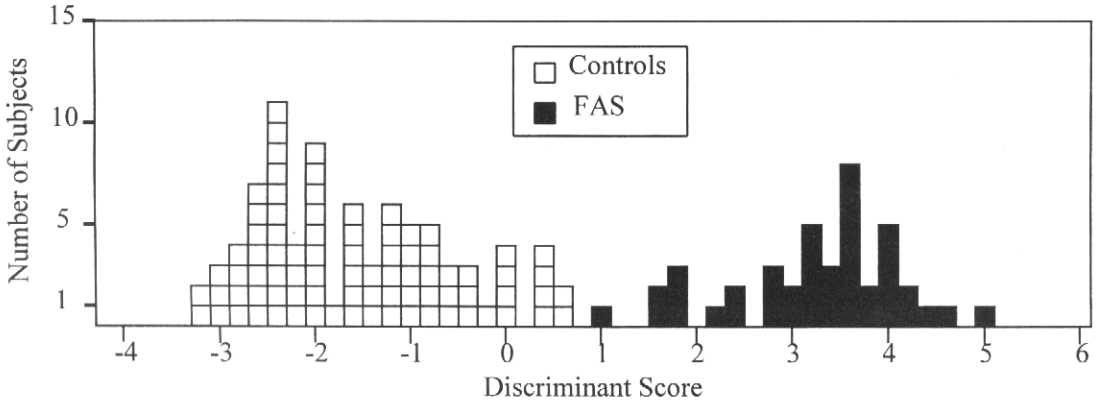


Fig. 4. Distribution of discriminant *(D)* scores among the patients with and without FAS, based on the following discriminant equation: D-score = 1.1075 – 6.0082 (palpebral fissure length/inner canthal distance ratio) + 1.1448 (philtrum Likert score) + 0.0066 (upper lip circularity). (The discriminant equation was derived from all 126 subjects.)

classifying a subject as FAS on the basis of screening (sensitivity = 100%, specificity = 100%, and overall accuracy = 100%).

Equation 3 reflects the most objective, sensitive, and specific discriminant function explaining 100% of the total variance (chi-square value [3, 126] = 226; $p = 0.0000$). The distribution of D-scores for all 126 subjects with and without FAS, derived from equation 3, is presented in Fig. 4.

Sensitivity and specificity across all six discriminant equations were unaffected by race, gender, and age in this study population.

Of the 84 control subjects, four had other syndrome diagnoses: Marfan syndrome, Williams syndrome, Dubowitz syndrome, and Aarskog syndrome. All discriminant equations correctly screened the four subjects as not having FAS.

DISCUSSION

This study has demonstrated that a phenotypic case definition of FAS can be derived from frontal facial photographs of individuals with FAS. It has also illustrated how this case-definition and methodologic approach can be used to develop an accurate and precise screening tool and diagnostic aid. The discriminating cluster of facial features identified from the photographs are identical to the facial features identified by direct facial measurement in a previous study,[7] thus supporting the validity of this photographic approach. Sensitivity and specificity were unaltered by race, gender, and age in both this and the previous study populations,[7] supporting the potential generalizability of this methodology. Further studies, however, will be necessary to evaluate more definitively the phenotypic variation that may occur across races.

It was particularly encouraging that the computer was able to identify a common phenotypic pattern in patients in whom FAS was diagnosed by more than one clinician, not only illustrating that a phenotypic consensus can be reached, but also illustrating a method by which to establish such a consensus. This is not to say that phenotypic classification does not vary across the clinical and research communities. One need only review the literature[8] and the published photographs to see how broadly the original[1, 9] and revised[10] descriptions of the FAS facial phenotype are interpreted, resulting in misclassification of research subjects and misdiagnosis of clinical patients. It is not that clinicians and researchers cannot consistently identify the facial phenotype; they simply do not have sufficiently specific guidelines for achieving such consistency.

To further test the ability of the function to differentiate between FAS and other syndromes with similar facial phenotypes, we obtained facial measures from the photographs of 10 additional control subjects published in syndrome diagnostic textbooks with the following syndromes: fetal hydantoin syndrome, Dubowitz syndrome, Noonan syndrome, Turner syndrome, Bloom syndrome, Aarskog syndrome, and Opitz syndrome. Ages ranged from 1 to 20 years, with a distribution comparable to that of the 126 subjects with and without FAS. All but the individual with fetal hydantoin syndrome were classified correctly as not having the facial phenotype of FAS. This suggests that this screening tool will effectively separate the dysmorphic features of FAS from normal and also from many, but not necessarily all, other syndromes.

A key factor in the success of this study was the quality of the photographic images. It is important to note that all photographs were taken by nonprofessional photographers using handheld cameras. Obtaining a quality photograph does not require sophisticated equipment or expertise. One need only focus on three elements: proper alignment,[11] proper exposure, and a relaxed facial expression, all of which can be easily attained by shooting a test roll of film with a properly aligned photograph in hand as a guide.

Table III. Contrasts in facial features and discriminant scores between study groups with and without FAS

Predictor and outcome variables	Control group (n = 84)	FAS group (n = 42)	Test statistic	*p*
Right palpebral fissure length/ inner canthal distance ratio				
Mean (SD)	0.83 (0.09)	0.67 (0.07)	$t = 10.1$	0.000
Minimum-maximum	0.65-1.12	0.60-0.88		
Philtrum, 5-point Likert scale: n(%)				
Deeply furrowed	47 (56.0)	0 (0)	U = 23.5*	0.0000
Somewhat furrowed	22 (26.2)	0 (0)		
Mid range	11 (13.0)	1 (2.3)		
Somewhat smooth	4 (4.8)	7 (16.7)		
Very smooth	0 (0)	34 (81.0)		
Philtrum luminosity†				
Mean (SD)	14.2 (9.0)	3.4 (2.5)	$t = 10.3$	0.000
Minimum-maximum	3-45	0-10		
Upper lip, 5-point Likert scale: n(%)				
Very thick	11 (13.0)	0 (0)	U = 509*	0.0000
Somewhat thick	13 (15.5)	1 (2.4)		
Mid range	21 (25.0)	1 (2.4)		
Somewhat thin	34 (40.5)	15 (35.7)		
Very thin	5 (6.0)	25 (59.5)		
Upper lip, circularity‡				
Mean (SD)	57.5 (14.8)	101.5 (49.3)	$t = -5.7$	0.000
Minimum-maximum	28.2-115.8	50.6-343.2		
Discriminant score, equation 1§				
Mean (SD)	−1.6 (1.1)	3.2 (0.8)	$t = -27.9$	0.000
Minimum-maximum	−3.9-0.7	1.1-4.4		
Discriminant score, equation 2#				
Mean (SD)	−1.6 (1.0)	3.2 (0.9)	$t = -25.6$	0.000
Minimum-maximum	−4.1-0.7	1.1-5.1		
Discriminant score, equation 3¶				
Mean (SD)	1.0 (1.0)	−1.9 (1.0)	$t = 15.0$	0.000
Minimum-maximum	−1.6-4.4	−5.7-−0.5		

*Mann-Whitney U test.
†Contrast in pixel luminosity between philtral ridge and furrow measured on a 256-unit gray scale.
‡Perimeter^2/area.
§Based on discriminant equation 1: D = 0.7408 − 5.7337 (palpebral fissure length/inner canthal distance ratio) + 1.1677 (philtrum Likert score) + 0.1587 (upper lip Likert score) derived across all 126 subjects.
#Based on discriminant equation 2: D = −6.4719 + 8.6104 (palpebral fissure length/inner canthal distance ratio) + 0.0767 (philtrum luminosity) − 0.0145 (upper lip circularity) derived across all 126 subjects.
¶Based on discriminant equation 3: D = 1.1075 − 6.0082 (palpebral fissure length/inner canthal distance ratio) + 1.1448 (philtrum Likert score) + 0.0066 (upper lip circularity) derived across all 126 subjects.

Recording facial measurements indirectly from a photograph rather than directly from the face has several advantages. By capturing a photographic or computer image, all three key facial features can be recorded on objective, continuous scales with the aid of computer software. This approach maximizes measurement accuracy and precision. If computer-derived continuous measures can be developed that validly capture the information obtained by the Likert scales, the process of phenotypic classification could be computer automated. The results of this study support the feasibility of such an endeavor. The continuous measure of circularity used to measure lip thinness was highly correlated with the Likert ranking (Spearman rank correlation coefficient 0.84: $p = 0.000$) and performed with equal sensitivity and specificity (100%). The continuous measure of luminosity used to measure philtrum smoothness was also highly correlated with the Likert ranking (Spearman rank correlation coefficient 0.74: $p = 0.000$). Although sensitivity and specificity dropped by 1% and 5%, respectively, modification of the luminosity measure could help regain the discriminating power. The increased accuracy, precision, and efficiency that could be achieved with this computerized photographic approach are all key to developing effective surveillance tools, screening tools, and diagnostic aids.

No two individuals with FAS necessarily have identical facial features; all, however, have the overall gestalt. To define the gestalt, one must take a multivariate approach, not a univariate checklist approach as illustrated in Fig. 3. The discriminant analysis used in this study accomplished this by identifying both the minimum number of features and the

magnitude of expression of each feature that most accurately differentiated individuals with and without the facial gestalt. The discriminant analysis identified short palpebral fissures, smooth philtrum, and thin upper lip as the minimum cluster of features needed to define the phenotype and differentiate individuals with highest accuracy. Further studies to confirm the validity and generalizability of these study results will be necessary. In the discriminant equation, the three facial features serve as the predictor variables, each with a beta coefficient reflecting their level of contribution to the overall gestalt appearance. The equation computes a discriminant score that, in essence, is a proxy measure of the gestalt appearance recorded on a continuous scale. To use this equation, one would measure the three facial features from a photograph, insert the values into the equation, compute the discriminant score, and classify the phenotype on the basis of whether the individual's discriminant score fell above or below the cutoff value. This equation not only provides a method by which to differentiate individuals with and without the facial phenotype of FAS but also provides a standardized objective language in which clinical researchers can describe the magnitude of expression of the phenotype in their study populations for comparative purposes.

Surveillance generally uses methods distinguished by their practicality, uniformity, and frequently their rapidity, rather than by complete accuracy.[12] To date, effective methods for FAS surveillance do not exist.[13] Passive surveillance such as hospital-based birth defects registries[14] examine only the medical record, not the patient, and focus on an age group when FAS is known to be misdiagnosed and underreported.[15] Passive surveillance has resulted in marked underestimation of FAS prevalence. In contrast, active surveillance relies on direct collection of data from patients and, although more costly, results in more valid and reliable data and in estimated rates of FAS that are an order of magnitude higher than those estimated from passive surveillance.[13] The results of this study demonstrate that computerized analysis of facial photographs has the potential for serving as a highly efficient, reproducible, and potentially highly accurate active surveillance tool. Photographic analysis has the following advantages: (1) photographic images are inexpensive, (2) professional expertise or sophisticated equipment is not needed for collection, (3) photographs can be stored and analyzed with complete anonymity by cropping the images, (4) data can be transferred electronically for centralized analysis, maximizing consistency of interpretation, (5) use of a highly accurate and objective case definition would give highly reproducible results with time, which is paramount to tracking trends, and (6) population-based surveillance of more representative segments of the population could be achieved because a broad age range can be accurately assessed and surveillance need not be restricted to hospital or research institutions, because of the ease with which data can be collected.

Primary prevention of secondary disabilities associated with FAS will require accurate and reliable screening tools and diagnostic aids. Unlike surveillance, screening and diagnosis target individuals for the purpose of early identification and intervention. An ideal screening tool is highly sensitive, specific, accurate, precise, reproducible, and valid. This tool has demonstrated all these qualities in this study population.

We express our thanks to Cliff Astley for his computer programming support, Erik McArthur for the computer capture of all photographic images, and Ann P. Streissguth, PhD, and Hans-Ludwig Spohr, MD, for their contribution of photographs to this study.

REFERENCES

1. Clarren SK, Smith DW. The fetal alcohol syndrome. N Engl J Med 1978;298:1063-7.
2. Abel EL, Sokol RJ. Incidence of fetal alcohol syndrome and economic impact of FAS-related anomalies. Drug Alcohol Depend 1987;19(1):51-70.
3. National Institute on Alcohol Abuse and Alcoholism. Seventh Special Report to the U.S. Congress on Alcohol and Health. Washington (DC): U.S. Department of Health and Human Services, 1990. DHHS publication No. ADM (90-1656).
4. Spohr HL, Willms J, Steinhausen HC. Prenatal alcohol exposure and long-term developmental consequences. Lancet 1993;341:907-10.
5. Aronson M, Olegard R. Children of alcoholic mothers. Pediatrician 1987;14(1-2):57-61.
6. Cordero JF, Floyd RL, Martin ML, Davis M, Hymbaugh K. Tracking the prevalence of FAS. Alcohol Health Research World 1994;18(1):82-5.
7. Astley SJ, Clarren SK. A fetal alcohol syndrome screening tool. Alcohol Clin Exp Res 1995;19:1565-71.
8. Ernhart CB, Greene T, Sokol RJ, Martier S, Boyd TA, Ager J. Neonatal diagnosis of fetal alcohol syndrome: not necessarily a hopeless prognosis. Alcohol Clin Exp Res 1995;19: 1550-7.
9. Jones KL, Smith DW. Recognition of fetal alcohol syndrome in early infancy. Lancet 1973;2:999-1001.
10. Sokol RJ, Clarren SK. Guidelines for use of terminology describing the impact of prenatal alcohol on the offspring. Alcohol Clin Exp Res 1989;13:597-8.
11. Farkas LG. Anthropometry of the head and face. 2nd ed. New York: Raven Press, 1994:79-88.
12. Last JM. A dictionary of epidemiology. 2nd ed. New York: Oxford University Press, 1988:141.
13. Stratton K, Howe C, Battaglia F, editors. Fetal alcohol syndrome: diagnosis, epidemiology, prevention, and treatment. Washington (DC): National Academy Press, 1995:82-99.
14. Centers for Disease Control and Prevention. Fetal alcohol syndrome—United States, 1979–1992. MMWR Morb Mortal Wkly Rep 1993;42:339-42.
15. Little BB, Snell LM, Rosenfeld CR, Gilstrap LCI, Gant NF. Failure to recognize fetal alcohol syndrome in newborn infants. Arch Pediatr Adolesc Med [Am J Dis Child] 1990;144: 1142-6.