Supplementary Online Content

May PA, Chambers CD, Kalberg WO, et al. Prevalence of fetal alcohol spectrum disorders in 4 US communities. *JAMA*. doi:10.1001/jama.2017.21896

This supplementary material has been provided by the authors to give readers additional information about their work.

eBox. Updated criteria for the classification of fetal alcohol spectrum disorders modified for use in the CoFASP consortium

I. Fetal Alcohol Syndrome

A classification of fetal alcohol syndrome requires all features specified in points A through D

A. A characteristic pattern of minor facial anomalies, including 2 or more of the following:

    1. Short palpebral fissures (≤10th percentile)

    2. Thin vermilion border of the upper lip (rank 4 or 5 on lip-philtrum guide)

    3. Smooth philtrum (rank 4 or 5 on lip-philtrum guide)

B. Growth deficiency

    1. Height, weight, or both at or below than the 10th percentile

C. Deficient brain growth

    1. Head circumference at or below the 10th percentile

D. Neurobehavioral impairmenta

    1. Point a, b, or both

        a. with cognitive impairment

        Evidence of global impairment (general conceptual ability ≥1.5 SD below the mean, or performance IQ or verbal IQ or spatial IQ ≥1.5 SD below the mean) or

        Cognitive deficit in at least 1 neurobehavioral domain of 1.5 or more SD below the mean (executive functioning, memory impairment, or visual-spatial impairment or 1.0 or more SD below the mean for specific learning impairment)

        b. with behavioral impairment without cognitive impairment

        Evidence of behavioral deficit in at least 1 domain 1.5 SD or more below the mean in impairments of self-regulation (mood or behavioral regulation impairment, attention deficit, or impulse control)

II. Partial Fetal Alcohol Syndrome

For children with documented prenatal alcohol exposure (see Section IV herein), a classification of partial fetal alcohol syndrome requires features A and B

A. A characteristic pattern of minor facial anomalies, including 2 more of the following:

    1. Short palpebral fissures (≤10th percentile)

    2. Thin vermilion border of the upper lip (rank 4 or 5 on lip-philtrum guide)

    3. Smooth philtrum (rank 4 or 5 on lip-philtrum guide)

B. Neurobehavioral impairment

    1. Same criteria as fetal alcohol syndrome

For children without documented prenatal alcohol exposure, a classification of partial fetal alcohol syndrome requires all features, points A through C

A. A characteristic pattern of minor facial anomalies, including 2 or more of the following:

    1. Short palpebral fissures (≤10th percentile)

    2. Thin vermilion border of the upper lip (rank 4 or 5 on lip-philtrum guide)

    3. Smooth philtrum (rank 4 or 5 on lip-philtrum guide)

B. Growth deficiency or deficient brain growth

    1. Height, weight or both at or below the 10th percentile, or

    2. Head circumference at or below the 10th percentile

C. Neurobehavioral impairment

    1. Same criteria as fetal alcohol syndrome

III. Alcohol-Related Neurodevelopmental Disorder

Requires features A and B

A. Documented prenatal alcohol exposure (see Section IV)

B. Neurobehavioral impairment (points a, b, or both)

    1. With cognitive impairment

      Evidence of global impairment (general conceptual ability ≥1.5 SD below the mean, or performance IQ, or verbal IQ, or spatial IQ ≥1.5 SD) or

      Cognitive deficit in at least 2 neurobehavioral domains of more than 1.5 SDs below the mean in executive functioning, memory impairment or visual spatial impairment or more than 1 SD below the mean for specific learning impairment

    2. With behavioral impairment without cognitive impairment

      Evidence of behavioral deficit in at least 2 domains 1.5 or more SDs below the mean in impairments of self-regulation (mood or behavioral regulation impairment, attention deficit, or impulse control)

IV. Alcohol Consumption Criteria

One or more of the following conditions must be met to constitute documented prenatal alcohol exposure during pregnancy. The information must be obtained from the biological mother or a reliable collateral source (eg, family member, social service agency, or medical record)

A. 6 or more drinks per week for 2 or more weeks during pregnancy

B. 3 or more drinks per occasion on 2 or more occasions during pregnancy

C. Documentation of alcohol-related social or legal problems in proximity to (prior to or during) the index pregnancy (eg, history of multiple citations for driving while intoxicated or history of treatment for an alcohol-related condition)

aModified from Hoyme et al, 2016[5]

Sample Characteristics

Demographic and other descriptors of the general population at the four participating sites as well as the same descriptors for the US population as a whole in 2015 are shown in eTable1. Sample sizes of each cohort, consent rates, and number of evaluations completed in each domain are described in eTable2.

Sampling Methods

A complete census (evaluation of all individuals in the selected study regions with respect to FASD), although theoretically ideal, was not feasible in the consortium, due in part to the extensive resources required to accomplish the needed child and maternal assessments. Instead, other methods for evaluating a sample of the population were employed, including random sampling in some settings and a screening approach in others. In the latter scenario, a larger sample (oversample) of children was screened on key criteria associated with prenatal alcohol exposure (i.e., growth or developmental concerns) and only those meeting the "high risk" screening criteria were selected for full evaluation. An additional sample of children who did not meet the screening criteria were also selected for full evaluation for purposes of creating control/comparison groups of children in each community. These comparison children represented the distribution of physical features, growth, and cognitive and behavioral functioning within each study community.

Three methods were used in the consortium sites for sampling. In some settings, the choice of method was due to practical issues. However, variability in the sampling techniques within or across sites provided the opportunity to evaluate the efficacy, practicality, and efficiency of different techniques of sampling. Each of the three methods are described below and outlined in the Figure in the main paper.

Sampling Method 1 (SM1): Oversample of Small Children with a Randomly-Selected Comparison Group from All Eligible Children

In this method, a tiered approach was used to most efficiently evaluate children in the study sample, and to provide a census or oversample of all small children. By oversampling the small children, the rationale was to increase the likelihood of identifying children with FAS or pFAS, as growth (weight, height or head circumference $\leq 10^{th}$ centile) are features that contribute to these two FASD classification categories. Permission was sought and obtained from parents/guardians for each child enrolled in first grade in all community schools at the site to participate in screening on weight, height and head circumference. Any child who met the criteria of $\leq 25^{th}$ centile on any one growth measure using the standard Centers for Disease Control and Prevention growth charts[1] was advanced to receive the physical examination by study dysmorphologists. Small children who had alcohol-related physical features identified in the dysmorphology examination went on to receive neurobehavioral evaluation and their mothers or collateral representatives were interviewed for maternal risk factors including extensive questions regarding alcohol consumption. In addition, randomly selected candidates for the comparison groups were drawn via random number generation programs from the entire pool of children enrolled in the first-grade classes of each school. All randomly-selected candidates who had been consented into the study were evaluated for all four relevant domains: growth; dysmorphology; neurobehavior; and maternal risk factors, especially drinking alcohol in pregnancy (see the Figure in the main paper).

In addition to screening on growth, case identification in SM1 was supplemented by teacher referrals of students who were not performing well in class, although the referrals were few to none at each SM1 site. In addition, children whose numbers were randomly selected from total class lists as candidates for the comparison groups and who were found to meet criteria for FASD were included in the count of FASD cases. This latter process enabled identification of those with an FASD irrespective of growth deficits, and provided proportions generated from random selection from the entire first grade class enrollment for weighting used to estimate prevalence. In the situation where twin pairs were eligible for the study and one member of the pair screened positive or was randomly selected, the other twin was also selected for the evaluation.

Sampling Method 2 (SM2): A Simple Random Sample

The second approach was a totally random sample drawn from all children enrolled in first grade within a particular community in a given year. There was no oversampling on growth, and there were no teacher referrals. Instead, full evaluations for growth, dysmorphology, neurobehavior, and maternal drinking were completed for all children chosen randomly from the entire enrollment of the first-grade classes at the study schools who were consented by parents or guardians to participate in the study. We first drew a sample (30 - 40%) from class lists of all enrolled children, explained the random nature of the sample to all parties in the community, and worked to gain consent for each selected child to participate in all aspects of the evaluation. Therefore, in SM2 all case and non-case children came from the random selection pool. Because there was no initial screening on growth, the full spectrum

of FASD including cases with no growth deficits and children with ARND, were likely to be evaluated with no census or oversample of small children.

Sampling Method 3 (SM3): Oversample of Small Children or Children with Developmental Concerns and a Random Sample of Children with Neither

In this approach, consented children from the selected study schools were first screened on growth using the same criteria as SM1 (i.e., weight, height or head circumference ≤25th centile). In addition, each consented child's parent or guardian completed the Parents' Evaluation of Developmental Status (PEDS) developmental assessment questionnaire.[2] All children who screened positive on growth and/or had two or more developmental concerns on the PEDS were selected to receive the full evaluation including the dysmorphology exam, neurobehavioral testing and assessment of maternal drinking. Consented children who had repeated the first grade in school were also considered screen positive and selected for the full evaluation, although this number was small. A random sample of consented children who screened negative on both growth and the PEDS developmental questionnaire was also selected for a comparison group, and received the full evaluations. Similar to SM1, children in the comparison group who met criteria for FASD were included in the case count for FASD and removed from the comparison group.

Dysmorphology Examination and FASD Categories

The dysmorphology evaluations for each child were conducted by one or more of a team of pediatricians at each site with specialized training and expertise in clinical genetics and dysmorphology and each was highly experienced in the diagnosis of FASD. Each exam was assisted on-site by a support team and scribe. The dysmorphologists used a standard examination form checklist including the cardinal facial features of FAS as well as multiple alcohol-associated minor anomalies (eFigure 1). Standard measurement equipment, reference charts for centiles,[1] palpebral fissure length,[3] inner canthal distance, outer canthal distance, and lip/philtrum guides[4] were used across sites.[5] As part of the physical evaluation, dysmorphologists also recorded the presence of Alcohol-Related Birth Defects (ARBD; e.g., heart or musculoskeletal defects) when such major structural anomalies were detected and additional information was available. However, ARBD as part of the FASD spectrum was not systematically evaluated as part of the CoFASP study and therefore is not included in the prevalence estimates. In addition, dysmorphologists documented alternative or suspected non-alcohol-related diagnoses, including genetic disorders, that may have ruled out FASD. Most evaluations were performed at the school sites, but in a small number of cases exams were performed at the study research office or the participant's home.

All examiners were blinded to the child's status regarding prenatal exposure to alcohol or neurobehavioral performance at the time of the physical exam. Two-dimensional (2D) facial images of children were captured at the time of the dysmorphology examination for purposes of assessing reliability of the qualitative assessment of features across multiple examiners. These images were used in the final case conferences held by the respective investigative teams to refresh memory of the dysmorphologists and examination team as findings for each child were discussed in preparation for the classification of FAS, pFAS, ARND, or not FASD.

Maternal or Collateral Interviews

A common core of maternal risk questions was agreed upon by all collaborators and the advisory committee early in the consortium initiative. These questions assessed maternal risk and protective factors via interviews performed in person at the Midwestern, Rocky Mountain, and Southeastern sites and over the telephone in almost all cases at the Pacific Southwestern site. Interviewers were blinded to the status of the child on dysmorphology and neurobehavior.

General maternal health and childbearing questions, demographic information, specific alcohol and other drug use by quantity, frequency, and gestational timing, and social or legal problems related to alcohol were addressed in each interview. The questionnaires were designed to allow for collection of information from a consented collateral source if the biological mother of the consented child was not available.

Interview protocols utilized questions on nutrition and diet, general health, and drinking items that were formulated, sequenced, and arranged in an overall health context via a timeline follow-back methodology.[6,7] The questions were designed to enhance recall and elicit accurate reporting of alcohol and other drugs consumed from a variety of sources and beverage types.[8,9]

Usual drinking patterns and drinking before pregnancy recognition were used to more accurately calibrate quantity and frequency of drinking during the index pregnancy.[10-15] The sequencing of questions had been developed in similar population-based studies and was used because direct reporting of prenatal drinking may be under-

reported in some settings.[16-19] Retrospective reports of maternal drinking have been found to reflect higher levels of consumption than those reported during the prenatal period.[15,19-22]

Each of the two research teams also included additional maternal risk questions that were unique to their respective settings. The design of the maternal questionnaires evolved from previous research on teratogens and FASD epidemiology.[12,23-26]

Cognitive and Behavioral Performance

Psychologists or psychometrists performed the neurobehavioral evaluations for consented children. They performed these evaluations either at the school site or research offices, and were blinded to the dysmorphology findings or maternal alcohol history for each child. The domains, test battery, and criteria for assessing cognitive performance and behavior with multiple tests and checklists were agreed upon early in the consortium initiative. The instruments used are presented in eFigure2 by domain of assessment and by cutoffs used to classify neurobehavioral deficits in this study. The battery was selected to evaluate the research-based deficits known to occur in children who were prenatally exposed to alcohol. In addition, the battery was chosen for cultural-relevance and appropriateness for the populations under study, and for availability in English and Spanish. Each measure was a standardized, proprietary test or checklist that is commonly used by assessment professionals to evaluate children in the first grade age range, and each item is accessible through standard media.

Case Conferences for Classification of Cases

Final classifications were made in case conferences where the findings for each child in each domain were discussed in a structured, sequential, roundtable fashion.[5,27] At each research team site, case conference participants included the site principal investigator and research team members who either performed or oversaw the dysmorphology examinations, neurobehavioral testing, or maternal interviews. While the findings were being presented and reviewed by the group, 2-dimensional digital photos of the child's face were projected to contextualize the data for the dysmorphologists and assessment team and to refresh memories of the dysmorphologist regarding the examination.

Consistency and Quality Assurance for the Dataset

In classifying children within the FASD continuum, consortium criteria were initially applied and later double-checked by the data managers for the research teams for consistency and accuracy. Classifications were then triple-checked by the consortium investigative teams by reciprocal exchange of all relevant data for all FASD cases and a sample of non-cases. Each team was blinded to the other team's classification for each case and was asked to determine whether the criteria had been applied accurately and consistently across sites.

Formulas for Calculation of the Prevalence by Sampling Method
Sampling Method 1 (SM1)

<u>Conservative Prevalence Estimate</u>

For FAS (and similarly for pFAS and ARND): $\left(\dfrac{X_{FAS}}{E}\right)k$

For Total FASD: $\left(\dfrac{X_{FAS}+X_{pFAS}+X_{ARND}}{E}\right)k$

<u>Weighted Prevalence Estimate</u>

For FAS (and similarly for pFAS): $\left(w_S\dfrac{X_{S:FAS}}{N_S}+w_{NS}\dfrac{X_{RSNS:FAS}}{N_{RSNS}}+w_{TR}\dfrac{X_{TR:FAS}}{N_{TR}}\right)k$

For ARND: $\left(w_{S_Q}\dfrac{X_{S_Q:ARND}}{N_{S_Q}}+w_{S_D}\dfrac{X_{SRS_D:ARND}}{N_{SRS_D}}+w_{NS}\dfrac{X_{RSNS:ARND}}{N_{RSNS}}+w_{TR}\dfrac{X_{TR:ARND}}{N_{TR}}\right)k$

For Total FASD:

$$\left(w_{S_Q}\dfrac{X_{S_Q:FAS}+X_{S_Q:pFAS}+X_{S_Q:ARND}}{N_{S_Q}}+w_{S_D}\dfrac{X_{SRS_D:ARND}}{N_{SRS_D}}+w_{NS}\dfrac{X_{RSNS:FAS}+X_{RSNS:pFAS}+X_{RSNS:ARND}}{N_{RSNS}}+\right.$$
$$\left. w_{TR}\dfrac{X_{TR:FAS}+X_{TR:pFASD}+X_{TR:ARND}}{N_{TR}}\right)k$$

Where:

Subsample labels are defined as follows:

- S: Children that were small on the initial screening.

- S_Q: Small children that had enough alcohol-related physical features on dysmorphology to meet criteria for FAS/pFAS.

- S_D: Small children that did not have enough alcohol-related physical features to qualify for FAS/pFAS.

- SRS_D: Small children selected in the random sample that did not have enough alcohol-related physical features to qualify for FAS/pFAS.

- NS: Children that were not small on the initial physical exam.

- $RSNS$: Children selected in the random sample that were not small on the initial physical exam.

- TR: Children who entered as a twin or referral.

E is the total number of children enrolled in all 1st grade classrooms at the study site.

M is the number of children that were consented and screened.

$M_{Subsample}$ is the number of children among the M total that belong to the subsample indicated by the subscript.

$N_{Subsample}$ is the number of children in the subsample indicated by the subscript that consented and were seen for a dysmorphology exam.

The weights w are the proportions of children enrolled in the study in each subsample. For the subsamples S, NS, and TR, $w_{Subsample} = \frac{M_{Subsample}}{M}$. For S_Q and S_D, $w_{S_Q} = w_S \frac{N_{S_Q}}{N_S}$ and $w_{S_D} = w_S \frac{N_{S_D}}{N_S}$.

X_{FAS}, X_{pFAS}, and X_{ARND} are the numbers of children classified with FAS, pFAS, and ARND, respectively.

$X_{Subsample: FAS}$, $X_{Subsample: pFAS}$, and $X_{Subsample: ARND}$ are the numbers of children in Subsample (one of S_Q, SRS$_D$, RSNS, or TR) classified with FAS, pFAS, and ARND, respectively.

k = constant (1,000).

Sampling Method 2 (SM2)

Sampling Method 2 was a simple random sample chosen from all children enrolled in first grade samples at a site. No pre-screening for growth or neurobehavioral concerns was done. Prevalence was calculated as follows:

Conservative and Alternative Prevalence Estimates

Conservative Prevalence Estimate for FAS (and similarly for pFAS and ARND): $\left(\frac{X_{FAS}}{M}\right) k$

Conservative Prevalence Estimate for Total FASD: $\left(\frac{X_{FAS}+X_{pFAS}+X_{ARND}}{M}\right) k$

Alternative Prevalence Estimate for FAS (and similarly for pFAS and ARND): $\left(\frac{X_{FAS}}{N}\right) k$

Alternative Prevalence Estimate for Total FASD: $\left(\frac{X_{FAS}+X_{pFAS}+X_{ARND}}{N}\right) k$

Where:

X_{FAS}, X_{pFAS}, and X_{ARND} are the numbers of children classified as FAS, pFAS and ARND.

M is the total number of randomly selected 1st grade children from school class rolls at the site.

N is the total number of randomly selected 1st grade children from study school class rolls at the site who were consented into the study and received full evaluations.

k = constant (1,000)

Sampling Method 3 (SM3)

Conservative Prevalence Estimate

Calculations are the same as in SM1.

Weighted Prevalence Estimate:

For FAS (and similarly for pFAS and ARND): $\left(w_{SP} \frac{X_{SP:FAS}}{N_{SP}} + w_{SN} \frac{X_{RC:FAS}}{N_{RC}} \right) k$

For Total FASD: $\left(w_{SP} \frac{X_{SP:FAS} + X_{SP:pFAS} + X_{SP:ARND}}{N_{SP}} + w_{SN} \frac{X_{RC:FAS} + X_{RC:pFAS} + X_{RC:ARND}}{N_{RC}} \right)$

Where:

Subsample labels are defined as follows:

- *SP*: Children that screened positive on the initial growth screening.

- *SN*: Children that screened negative on the initial growth screening.

- *RC*: Children selected in the random sample.

M is the number of children that were consented and screened.

M_{SP} and M_{SN} are the numbers of children among the *M* total that belong to the subsamples SP and SN, respectively.

N_{SP} and N_{RC} are the numbers of children with any classification (FAS, pFAS, ARND or No FASD) in the subsamples SP and RC.

$w_{SP} = \frac{M_{SP}}{M}, w_{SN} = \frac{M_{SN}}{M}$.

$X_{Subsample: FAS}$, $X_{Subsample: pFAS}$, and $X_{Subsample: ARND}$ are the numbers of children in Subsample (one of SP or RC) classified with FAS, pFAS, and ARND, respectively.

k = constant (1,000).

Variance Estimates and Confidence Intervals

To obtain confidence intervals for the prevalence estimates, the variance of the estimated prevalence was estimated by $\text{var}(\hat{p}) = \sum_{j=1}^{J} w_j^2 \, \text{var}(\hat{p}_j)$, where $\hat{p}_j = X_j / N_j$. For SM1 and SM2, the subsample variances are estimated by $\text{var}(\hat{p}_j) = \hat{p}_j (1 - \hat{p}_j) / N_j$. For SM3, cluster sampling (where cluster = school) was accounted for using nonparametric bootstrap, by resampling with replacement the clusters (i.e., schools) with 10,000 bootstrap runs. Bootstrap sampling was performed independently in the Screen Positive and Random Control subsamples in each sample. The prevalence was computed in each bootstrap sample, and the variance of the 10,000 prevalence statistics was computed to estimate $\text{var}(\hat{p}_{SP})$ and $\text{var}(\hat{p}_{RC})$. For all prevalence estimates, to ensure that the confidence interval bounds were between 0 and 1, the transformation $g(x) = \log(-\log(x))$ was first applied. Normal confidence bounds of form $g(\hat{p}) \pm z \cdot \text{se}(g(\hat{p}))$ where computed, where z is the 0.975 quantile of a standard normal distribution, $\text{se}(g(\hat{p})) = \sqrt{\widehat{\text{var}(g(\hat{p}))}}$, and $\widehat{\text{var}(g(\hat{p}))} \approx \text{var}(\hat{p}) / (\hat{p} \log(\hat{p}))^2$ by the delta method. The resulting interval was then back-transformed to the original scale. In Sample 2 for the Midwestern City there were no cases of FAS; here the confidence interval was computed as (0, 3/N) using the "Rule of 3" method, where N = 236 was the total number of children with a classification.

References

1. Kuczmarski RJ, Ogden CL, Guo SS, et al. CDC growth charts for the United States: methods and development. *Vital Health Stat*. 2000 May; 1-190.

2. Woolfenden S, Eapen V, Williams K, Hayen A, Spencer N, Kemp, L. A systematic review of the prevalence of parental concerns measured by the Parents' Evaluation of Developmental Status (PEDS) indicating developmental risk.*BMC Pediatrc*. 2011;13(14):231.

3. Thomas T, Galtantzis YA, Frias JL.Palpebral fissure length from 29 weeks gestation to 14 years.*J Pediatr*. 1987;111(2):267-268.

4. Astley SJ, Clarren, S. Diagnosing the full spectrum of fetal alcohol-exposed individuals: introducing the 4-digit diagnostic code. *Alcohol Alcohol*. 2000;35(4):400-10.

5. Hoyme HE, Kalberg WO, Elliott AJ, et al. Updated clinical guidelines for diagnosing fetal alcohol spectrum disorders. *Pediatr*.2016;138(2).

6. Sobell LC, Sobell MB, Leo Gl, Cancilla A.Reliability of a timeline method: assessing normal drinker's reports of recent drinking and a comparative evaluation across several populations. *Br J Addict*. 1988;83: 393-402.

7. Sobell LC, Agrawal S, Annis H, et al. Cross-cultural evaluation of two drinking assessment instruments: alcohol timeline followback and inventory of drinking situations. *Subst Use Misuse*. 2000;36: 313-331.

8. Jacobson SW, Jacobson JL, Sokol RJ, Martier SS, Ager JW, Kaplan MG. Maternal recall of alcohol, cocaine, and marijuana use during pregnancy. *Neurotoxicol Teratol*. 1991;13: 535-540.

9. King AC.Enhancing the self-report of alcohol consumption in the community: two questionnaire formats. *Am J Pub Health*. 1994;84: 294-296.

10. May PA, Brooke L, Gossage JP, et al. Epidemiology of fetal alcohol syndrome in a South African community in the Western Cape Province. *Am J Pub Health*. 2000; 90(12): 1905-1912.

11. May PA, Gossage JP, Marais AS, et al.The epidemiology of fetal alcohol syndrome and partial FAS in a South African community. *Drug Alcohol Depend*. 2007;88(2-3): 259-271.

12. May PA, Gossage JP, Marais AS, et al. Maternal risk factors for fetal alcohol syndrome and partial fetal alcohol syndrome in South Africa: a third study. *Alcohol Clin Exp Res*. 2008;32(5):738-753. doi:10.1111/j.15300277.2008.00634.x.

13. May PA, Blankenship, Marais AS, et al. Maternal alcohol consumption producing fetal alcohol spectrum disorders (FASD): Quantity, frequency, and timing of drinking. *Drug Alcohol Depend*. 2013;133(2):502-512.

14. Viljoen DL, Croxford J, Gossage JP, Kodituwakku PW, May PA. Characteristics of mothers of children with Fetal Alcohol Syndrome in the Western Cape Province of South Africa: a case control study. *J Stud Alcoho*. 2002; 63: 6-17.

15. Czarnecki OM, Russell M, Cooper ML, Salter D.Five-year reliability of self-reported alcohol consumption. *J Stud Alcohol*. 1990;51: 68-76.

16. Morini L, Marchei E, Tarani L, et al. Testing ethylglucuronide in maternal hair and nails for the assessment of fetal exposure to alcohol: comparison with meconium testing. *Ther Drug Monit*. 2013;35: 402-407.

17. Wurst FM, Kelso E, Weinmann W, Pragst F, Yegles M, Sundström Poromaa I.Measurement of direct ethanol metabolites suggests higher rate of alcohol use among pregnant women than found with the AUDIT—a pilot study in a population-based sample of Swedish women. *Am J Obstet Gynecol*. 2008;198: 407.e1-407.e5.

18. Pichini S, Marchei E, Vagnarelli F, et al. Assessment of prenatal exposure to ethanol by meconium analysis: results of an Italian multicenter study. *Alcohol Clin Exp Res*. 2012;36: 417-724.

19. Alvik A, Haldorsen T, Groholt B, Lindemann R. Alcohol consumption before and during pregnancy: comparing concurrent and retrospective reports. *Alcohol Clin Exp Res*. 2006;30: 510-515.

20. Hannigan JH, Chiodo LM, Sokol RJ, et al. A 14-Year retrospective maternal report of alcohol consumption in pregnancy predicts pregnancy and teen outcomes. *Alcohol*. 2010;44: 583-594.

21. Jacobson JL, Jacobson SW, Sokol RJ, Martier SS, Ager JW, Shankaran S.Effects of alcohol use, smoking, and illicit drug use on fetal growth in black infants. *J Pediatr*. 1994;124: 757-764.

22. Jacobson SW, Chiodo LM, Sokol RJ, Jacobson JL. Validity of maternal report of prenatal alcohol, cocaine, and smoking in relation to neurobehavioral outcome. *Pediatr*. 2002; 109: 815-825.

23. Ceccanti M, Fiorentino D, Coriale G, et al. Maternal risk factors for fetal alcohol spectrum disorders in a province in Italy. *Drug Alcohol Depend*. 2014;145:201-2078.

24. Chambers CD, Yevtushok L, Zymak-Zakutnya N, et al. Prevalence and predictors of maternal alcohol consumption in 2 regions of Ukraine. *Alcohol Clin Exp Res*. 2014;38(4):1012-1019.
25. May PA, Gossage JP, Brooke LE et al.Maternal risk factors for fetal alcohol syndrome in the Western Cape Province of South Africa: a population-based study. *Am J Pub Health*. 2005;95(7):1190-1199.
26. May PA, Tabachnick BG, Gossage JP. Maternal factors predicting cognitive and behavioral characteristics of children with fetal alcohol spectrum disorders. *J Deve Behav Pediatr*. 2013;34(5): 314-325.
27. Hoyme HE, May PA, Kalberg WO, et al. A practical clinical approach to diagnosis of fetal alcohol spectrum disorders: clarification of the 1996 Institute of Medicine criteria. *Pediatr*. 2005; 115(1): 39-47.

eTable1. Demographic Indicators by Consortium Site and US

Demographic Indicator	Midwestern City	Rocky Mountain City	Southeastern County	Pacific Southwest City	United States
Population (7/2015)[1] (percentage of US population)	171,544 (0.05%)	59,638 (0.02%)	206,392 (0.06%)	1,406,630 (0.44%)	321,418,820 (100%)
Population change (%) since 2010[1]	11.4%	0.9%	9.4%	8.1%	4.1%
Race/Hispanic Ethnicity (2010)[1]					
White, non-Hispanic	84.9%	86.7%	68.2%	58.9%	63.7%
Black, non-Hispanic	4.2%	1.1%	17.8%	6.7%	12.6%
American Indian and Alaskan Native	2.7%	5.0%	0.4%	0.6%	0.9%
Asian	1.8%	0.9%	1.6%	15.9%	4.8%
Two or more races	2.5%	3.8%	2.3%	5.1%	2.9%
Hispanic or Latino	4.4%	3.4%	10.8%	28.8%	16.3%
Foreign born persons[1]	7.1%	2.2%	7.5%	26.6%	13.1%
Age – years (median)	34.5	38.9	36.4	33.6	37.2
Housing[1] Median household value	$155,200	$158,900	$167,700	$463,300	$176,700
Education[1] High School graduate or higher, % ages ≥25 years	90.8%	91.1%	84.6%	87.3%	86.3%
Bachelor's degree or higher, % ages ≥25 years	32.5%	25.5%	22.1%	43.0%	29.3%
Economy[1] Per capita income in past 12 months (2014 dollars)	$28,120	$24,733	$25,544	$33,902	$28,555
Median household income	$52,607	$43,374	$47,694	$66,116	$53,482
Persons in poverty	11.8%	16.1%	12.2%	15.4%	14.8%
Health Behavior Overall state health Rank in US[2]	15-19	20-25	30-34	15-19	Median 25 (Range 1-50)
Alcohol Use Binge drinking^ state %, (US rank)[2]	17.4% (35)	18.9% (41)	13.6% (9)	15.6%(21)	16.8% (25)
Excessive drinking[+], state % (US rank)[2]	18.3% (30)	20.8% (42)	15.1% (9)	17.2% (22)	Median = 17.4% (25)
Excessive drinking, county[3]	18.0%	20.0%	16.0%	20.0%	
Heavy drinking[#,] city[3]	6.4%	4.9%	4.9%	5.7%	Mean = 16.8%
State per capita ethanol consumption (2009), volume per person 14 years and older[4]	2.62 gallons 9.91 liters	2.99 gallons 11.32 liters	2.02 gallons 7.65 liters	2.33 gallons 8.82 liters	2.30 gallons 8.71 liters

Sources:
1. US Census

2. United Health Foundation, America's Health Rankings, 2015; comprised of scores on behaviors, community and environment, policy and clinical care; scores are ranked for each of the 50 states with better scores resulting in a higher rank among the 50 states; ranges indicate that different rankings are provided for each of the four domains named above
3. BRFSS (Behavioral Risk Factory Survey data of the CDC. Reported in local city and county statistical reports
4. La Valle and Yi, NIAAA Surveillance Report #92

^Binge drinking defined as: during the past 30 days, the consumption of 5 or more drinks for men or 4 or more drinks for females on an occasion

#Heavy drinking is defined as males having more than two drinks per day and females having more than one drink per day

+Excessive drinking of alcohol is defined as both binge drinking (above) and chronic drinking also referred to as heavy drinking (above)

eTable2. Detailed Characteristics of the Consortium Sites and Samples

Sample Number	Sample Method (SM)	Year Initiated	Growth &/or Development Screening Criteria	Number Elementary Schools in City or County	Number Schools in Study	Number Students Enrolled in 1st Grade	Agreed to Contact N(%)	Consent to Screen N(%)	Children Screened N(%)	Eligible for Full Evaluation N(%)	Consent to Full Evaluation N(%)	Dysmorphology Exam N(%)	Maternal Interview N(%)	Neurodevelopmental Testing N(%)
Sample 1: Midwest City	SM1	2010	($\leq$25th)^a	32	32	2,033	n/a	1,433 (70.5%)	1,433 (100.0%)	Screen Positive 318 (22.1%); Random Control 194 (13.5%)	512 (100.0%)	512 (100.0%)	153/287[f] (53.3%)	187/287[f] (65.1%)
Sample 2: Midwest City	SM2	2012	None			2,014	n/a	n/a	Random Sample[d]	709 (35.2%)	380 (53.6%)	379 (99.7%)	227 (59.7%)	236 (62.1%)
Sample 3: Rocky Mountain City	SM1	2012	($\leq$25th)^a	17	17	915	n/a	571 (62.4%)	561 (98.2%)	Screen Positive 239 (42.6%); Random Control 82 (14.6%)	321 (100.0%)	265 (82.6%)	126/168[f] (75.0%)	167/168[f] (99.4%)
Sample 4: Rocky Mountain City	SM2	2013	None			888	n/a	n/a	Random Sample[d]	400 (45.0%)	208 (52.0%)	206 (99.0%)	140 (67.3%)	203 (97.6%)
Sample 5: Southeastern County	SM1	2013	($\leq$25th)^a	24	14[c]	1,339	n/a	1,239 (92.5%)	1,217 (98.2%)	Screen Positive 262 (21.5%); Random Control 140 (11.5%)	402 (100.0%)	382 (95.0%)	196/284[f] (69.0%)	220/284[f] (77.5%)

Sample Number	Sample Method (SM)	Year Initiated	Growth &/or Development Screening Criteria	Number Elementary Schools in City or County	Number Schools in Study	Number Students Enrolled in 1st Grade	Agreed to Contact N(%)	Consent to Screen N(%)	Children Screened N(%)	Eligible for Full Evaluation N(%)	Consent to Full Evaluation N(%)	Dysmorphology Exam N(%)	Maternal Interview N(%)	Neurodevelopmental Testing N(%)
Sample 6: Southeastern County	SM1	2014	($\leq$25th,[a])			1,548	n/a	1,345 (86.9%)	1,341 (99.7%)	Screen Positive 294 (64.9%) / Random Control 159 (35.1%)	453 (100.0%)	443 (97.8%)	184/316[f] (58.2%)	249/316[f] (78.8%)
Sample 7: Pacific Southwestern City	SM3	2012	($\leq$25th,[a]) or PEDS[b]	201	27	2,238	1,100 (49.2%)[e]	831 (37.1%)[e]	704 (84.7%)[e]	Screen Positive 464 (65.9%) / Random Control 103 (14.6%)	427 (75.3%)[e]	409 (95.8%)	392 (91.8%)	419 (98.1%)
Sample 8: Pacific Southwestern City	SM3	2013	($\leq$25th,[a]) or PEDS[b]			2,171	1,027 (47.3%)[e]	802 (36.9%)[e]	798 (99.5%)[e]	Screen Positive 515 (64.5%) / Random Control 119 (14.9%)	499 (78.7%)[e]	487 (97.6%)	480 (96.2%)	492 (98.6%)
GRAND TOTAL	--	--	--	**219**	**90**	**13,146**		**6,809/ 11,358 (59.9%)[g]**		**3,998**	**3,202**	**3,083**	**1,898**	**2,173**

[a] $\leq$25th centile on height, weight, or head circumference

[b] Two or more concerns reported on the Parents' Evaluation of Developmental Status (PEDS), a screening tool that assesses language, motor, self-help, early academic skills, behavior and social-emotional/mental health.

[c] In the Southeastern County, there were 2 independent public school districts. 1 had 5 elementary schools, and all 5 were included in the study. The other district had 19 elementary schools, and 9 were chosen via random selection to participate in the study

[d] Eligible children in SM2 samples were defined as those selected completely random; there was no screening for growth. Parents or guardians of all eligible (i.e., randomly selected) children were asked to consent to have their children screened with the dysmorphology exam

eFigure1: Dysmorphology Checklist

Child's Name: _____   Child ID: _____

Exam Date (mm/dd/yyyy): ____ / ____ / _____   Name of Scribe: _____

Date of Birth (mm/dd/yyyy): ____ / ____ / _____   Name of Examiner: _____

Current Age: _____ Years _____ Months   Examination Site: _____

Child Sex: Child Race/Ethnicity:

☐ Male ☐ Female   ☐ Black ☐ Asian ☐ Hispanic ☐ Native American ☐ White ☐ Other (_____)

		Significant Observation?			**Percentile**
1. OFC	_____.__ (cm)	≤10%	Yes ☐	No ☐	OFC Percentile _____
2. Height	_____.__ (cm)	≤10%	Yes ☐	No ☐	Height Percentile _____
3. Weight	_____.__ (kg)	≤10%	Yes ☐	No ☐	Weight Percentile _____
4. Child's BMI: _____ Normal BMI for Age: _____					BMI Percentile _____
Head/Face		**Significant Observation?**			**Percentile**
5. ICD	_____.__ (cm)	≤25%	Yes ☐	No ☐	ICD Percentile _____
6. IPD	_____.__ (cm)	≤25%	Yes ☐	No ☐	IPD Percentile _____
7. PFL Left	_____.__ (cm)	≤10%	Yes ☐	No ☐	PFL-Left Percentile _____
8. PFL Right	_____.__ (cm)	≤10%	Yes ☐	No ☐	PFL-Right Percentile _____
9. OCD	_____.__ (cm)				OCD Percentile _____
10. Maxillary arc	_____.__ (cm)				
11. Mandibular arc	_____.__ (cm)				
12. Philtrum Length	_____.__ (cm)	Yes ☐		No ☐	Philtrum Percentile _____
13. Philtrum Lipometer Code	____ (1 – 5)	Yes ☐		No ☐	
14. Vermilion Lipometer Code	____ (1 – 5)	Yes ☐		No ☐	
15. Prognathism		Yes ☐		No ☐	
16. Midface Hypoplasia		Yes ☐		No ☐	
Ears		**Significant Observation?**			**Unilateral or Bilateral?**
17. RR track ears		Yes ☐		No ☐	Unilateral ☐ Bilateral ☐
18. Cupped		Yes ☐		No ☐	Unilateral ☐ Bilateral ☐
19. Low-Set		Yes ☐		No ☐	Unilateral ☐ Bilateral ☐
Eyes		**Significant Observation?**			**Unilateral or Bilateral?**
20. Strabismus		Yes ☐		No ☐	Unilateral ☐ Bilateral ☐
21. Ptosis		Yes ☐		No ☐	Unilateral ☐ Bilateral ☐
22. Epicanthal folds		Yes ☐		No ☐	Unilateral ☐ Bilateral ☐
Nose		**Significant Observation?**			
23. Flat nasal bridge		Yes ☐		No ☐	
24. Anteverted nose		Yes ☐		No ☐	
Hands		**Significant Observation?**			**Unilateral or Bilateral?**
25. Hypoplastic nails		Yes ☐		No ☐	Unilateral ☐ Bilateral ☐
26. 5th finger clinodactyly		Yes ☐		No ☐	Unilateral ☐ Bilateral ☐
27. Camptodactyly		Yes ☐		No ☐	Unilateral ☐ Bilateral ☐

Hand Creases	Yes ☐ No ☐		Unilateral or Bilateral?	
28. Hockey stick	Yes ☐	No ☐	Unilateral ☐	Bilateral ☐
29. Single transverse	Yes ☐	No ☐	Unilateral ☐	Bilateral ☐
30. Hypoplastic thenar	Yes ☐	No ☐	Unilateral ☐	Bilateral ☐
31. Other aberrant crease	Yes ☐	No ☐	Unilateral ☐	Bilateral ☐
Arms	**Significant Observation?**		**Unilateral or Bilateral?**	
32. Decreased pronation/supination	Yes ☐	No ☐	Unilateral ☐	Bilateral ☐
Hips/Legs/Feet	**Significant Observation?**		**Unilateral or Bilateral?**	
33. Knee contractures	Yes ☐	No ☐	Unilateral ☐	Bilateral ☐
34. Other contractures	Yes ☐	No ☐	Unilateral ☐	Bilateral ☐
35. Hip contractures	Yes ☐	No ☐	Unilateral ☐	Bilateral ☐
Other Body	**Significant Observation?**			
36. Hirsute	Yes ☐	No ☐		
37. Heart Murmur	Yes ☐	No ☐		
38. **Heart Defect**	**Yes ☐**	**No ☐**		
39. Ventricular septal defect	Yes ☐	No ☐		
40. Atrial septal defect	Yes ☐	No ☐		
41. Other heart defect	Yes ☐	No ☐	Specify: _____	
42. **Mental Status/Behavior**	**Yes ☐**	**No ☐**		
43. Hyperactive	Yes ☐	No ☐		
44. **Neurological**	**Yes ☐**	**No ☐**		
45. Hypertonic	Yes ☐	No ☐		
46. Hypotonic	Yes ☐	No ☐		
47. Seizures	Yes ☐	No ☐		
Other Measurements or Comments:				
Other Non-FASD Diagnosis:				

OFC: Occipital Frontal Circumference; BMI: Body Mass Index; ICD: Inner Canthal Distance; IPD: Interpupillary Ddistance; PFL: Palpebral Fissure Length; OCD: Outer Canthal Distance; RR track ears: Railroad Track Ears

eFigure2. Neurobehavioral Testing Battery and Cut-Off Criteria

Domains Evaluated	Types of Measures	Tools Used	Cut-off Criteria
Cognitive	General Intelligence	Differential Ability Scales - DASII	**Cognitive** DASII: Standard Score ≤ 79; (≥1.5 SD); percentile ≤ 8 NEPSY: Scaled Score ≤ 6; (≥1.5 SD); percentile ≤ 8 VMI: Standard Score ≤ 79; (≥1.5 SD); percentile ≤ 8
	Neurobehavioral Abilities • Executive functioning • Memory • Visual spatial	NEPSY • Speeded Naming subtest • Inhibition subtest • Visuomotor Precision subtest VMI (Visual-Motor Integration)	
Academic Achievement	Learning • Math • Reading • Spelling	Bracken Basic Concepts Scale	**Academic Achievement** Standard Score ≤ 85; Scaled Score ≤ 7; (≥1.0 SD); percentile ≤ 16
Behavior	• Mood or behavior regulation • Attention • Impulse control • Spelling	Achenbach Child Behavior Checklist (CBCL) – Parent Achenbach Teacher Report Form (TRF) – Teacher	**Behavior** T-Score ≥ 64; (≥1.5 SD); percentile ≥ 92
Adaptive Skills	• Daily Living • Communication • Socialization • Motor	Vineland Adaptive Behavior Scales	**Adaptive Skills** Standard Score ≤ 79; (≥1.5 SD); percentile ≤ 8

Domains Evaluated = neurobehavioral domains that were assessed
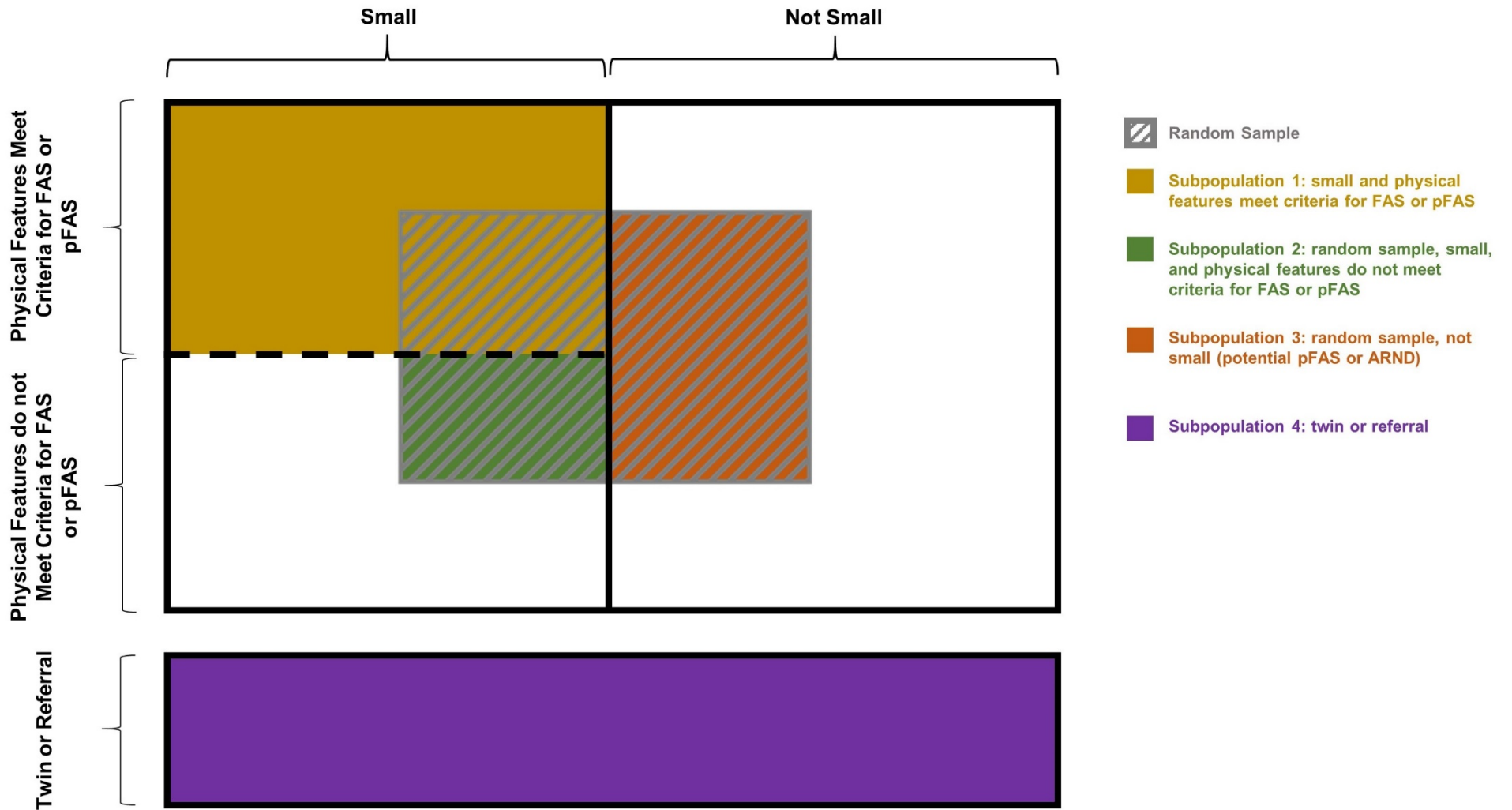Types of Measures = the areas of functioning within each domain that were assessed
Tools Used = the tools or tests used to assess the named area of functioning
Cut-off Criteria = cut-offs were used to define the scores indicating impairment for the specified domains on each test
Within the Cognitive Domain, arrows are used to align the types of measures assessed to the specific test/subtests used to evaluate each ability (e.g., the NEPSY Inhibition subtest was used to assess both executive functioning and memory, while both the NEPSY Visuomotor Precision subtest and the VMI were used to assess visual spatial abilities)
Abbreviations: NEPSY: A Developmental Neuropsychological Assessment; VMI: Beery-Buktenica Developmental Test of Visual-Motor Integration; CBCL: Child Behavior Checklist; TRF: Teacher Report Form

eFigure3. Subpopulations in SM1 Coded by Color

Abbreviations: SM: Sampling Method; FAS: Fetal Alcohol Syndrome; pFAS: partial Fetal Alcohol Syndrome; ARND: Alcohol-related Neurodevelopmental Disorder. "Small" is defined as children that are ≤ 25th centile for height, weight, or head circumference. Below we give an example calculation of the estimated prevalence of ARND in the Midwestern City for SM1. The color-coded boxes in the figure represent the four subpopulations: screened small with alcohol-related physical features meeting criteria for FAS or pFAS (gold), screened small and randomly selected but insufficient number of alcohol related physical features to meet criteria for FAS or pFAS (green), randomly selected and not small (orange); and twin or teacher referral (purple). For each subpopulation, a prevalence term X_j/N_j is

computed from children in the subpopulation that have a full evaluation. The four prevalences are then weighted according to the proportion of consented and screened children in each subpopulation and are summed as follows:

$$\left(w_{SQ} \frac{X_{SQ:ARND}}{N_{SQ}} + w_{SD} \frac{X_{SRS_D:ARND}}{N_{SRS_D}} + w_{NS} \frac{X_{RSNS:ARND}}{N_{RSNS}} + w_{TR} \frac{X_{TR:ARND}}{N_{TR}} \right) k$$

$$= \left(w_S \frac{N_{SQ}}{N_S} \frac{X_{SQ:ARND}}{N_{SQ}} + w_S \frac{N_{SD}}{N_S} \frac{X_{SRS_D:ARND}}{N_{SRS_D}} + w_{NS} \frac{X_{RSNS:ARND}}{N_{RSNS}} + w_{TR} \frac{X_{TR:ARND}}{N_{TR}} \right) k$$

$$= \left(\frac{M_S}{M} \frac{N_{SQ}}{N_S} \frac{X_{SQ:ARND}}{N_{SQ}} + \frac{M_S}{M} \frac{N_{SD}}{N_S} \frac{X_{SRS_D:ARND}}{N_{SRS_D}} + \frac{M_{NS}}{M} \frac{X_{RSNS:ARND}}{N_{RSNS}} + \frac{M_{TR}}{M} \frac{X_{TR:ARND}}{N_{TR}} \right) k$$

$$= \left(\frac{397}{1433} \frac{191}{390} \frac{X_{SQ:ARND}}{N_{SQ}} + \frac{397}{1433} \frac{199}{390} \frac{X_{SRS_D:ARND}}{N_{SRS_D}} + \frac{1036}{1433} \frac{X_{RSNS:ARND}}{N_{RSNS}} + \frac{3}{1433} \frac{X_{TR:ARND}}{N_{TR}} \right) k$$

$$= \left(\frac{397}{1433} \frac{191}{390} \frac{2}{191} + \frac{397}{1433} \frac{199}{390} \frac{1}{65} + \frac{1036}{1433} \frac{1}{119} + \frac{3}{1433} \frac{0}{3} \right) 1000$$

$$= 9.670809.$$