certificates that reported these defects in 76,863 babies born during 1989 and 1990. Compared with MACDP, birth certificates had a sensitivity of 19% for these defects, ranging from 38% for cleft lip/palate to 6% for malformed genitalia. The PVP was 78%, ranging from 100% (neural tube defects, rectal atresia/stenosis, esophageal atresia, and omphalocele/gastroschisis) to 60% for diaphragmatic hernia. Even for defects recognizable at birth, rates based on data from the revised birth certificates underestimated true birth defect rates. Birth certificates should be used cautiously as the sole ascertainment source for surveillance and epidemiologic studies.

Population-Based Case-Control Study of Gastroschisis: Reproductive and Demographic Factors and Substance Use. C. P. Torfs,* E. M. Velie, and F. W. Oechsli (California Birth Defects Monitoring Program, Emeryville, CA 94608).

Gastroschisis is a severe birth defect with a prevalence rate of about 2 per 10,000 births. The authors conducted a case-control study in the California Birth Defects Monitoring Program surveillance area. Cases were the first 110 infants with gastroschisis born after March 1988 whose mothers were Caucasian or Caucasian/Amerindian. For each case, two age-matched controls of the same ethnicity without a birth defect were randomly sampled from the same area. An in-home interview was administered to all mothers. This analysis covers maternal reproductive and parental demographic factors as well as substance abuse. The overall mean maternal age was 22.9 years. Univariate matched pairs analysis gave significant or borderline significant odds ratios for incomplete high school education versus none (odds ratio (OR) = 4.1, 95% confidence interval (CI) 1.5–11.1); one or more elective terminations (OR = 1.96, 95% CI 1.2–3.2); one to two packs of cigarettes (OR = 1.8, 95% CI 0.9–3.4); cocaine use (OR = 3.5; 95% CI 1.3–9.5); amphetamine use (OR = 2.6, 95% CI 1.3–5.3); marijuana use (OR = 2.8, 95% CI 1.6–5.4); and use of lysergic acid diethylamide (LSD) (OR = 10.0, 95% CI 1.2–86.4). Maternal and paternal drug use were highly correlated. When conditional logistic regression was used, the final set of variables was associated with 17% of the total deviance. Of the recreational drugs, alcohol use explained more than cigarette smoking, and cocaine contributed the most to a reduction in deviance.

Influence of Maternal Smoking, Paternal Smoking, and Involuntary Maternal Smoke Exposures on Oral Cleft Defects. G. M. Shaw* and C. R. Wasserman (California Birth Defects Monitoring Program, Emeryville, CA 94608).

Investigations have examined whether women who smoke during early pregnancy are at increased risk to have an infant with an oral cleft. Although results have been mixed, no study has estimated risks from both maternal and paternal smoking, nor from involuntary smoke exposures of women during pregnancy. To investigate these exposures, data were examined from a population-based case-control study of 555 oral cleft cases and 652 random liveborn controls from selected areas of California. Telephone interviews were conducted with the mothers of 460 cases and 482 controls. The unadjusted odds ratios for maternal smoking from

1 month preceding conception through the first trimester were 1.4 (95% confidence interval (CI) 1.0–1.9) and 2.2 (95% CI 1.3–3.6) for 1–19 and ≥20 cigarettes/day, respectively. For paternal smoking any time from 3 months preceding conception through the first trimester, the odds ratios were 1.1 (95% CI 0.82–1.6) and 1.7 (95% CI 1.1–2.5) for 1–19 and ≥20 cigarettes/day. The odds ratios for cleft lip with or without cleft palate were similar to those for cleft palate alone. Among many involuntary smoke exposures, the odds ratio for women who reported exposure to others' cigarette smoke in the workplace was 1.7 (95% CI 1.2–2.4). Adjusted estimates for combinations of maternal and paternal smoke exposures, for other involuntary exposures, and for isolated and multiple case groups will be presented.

Monitoring Study of Congenital Malformations of the Central Nervous System. B. Wang,* F. Li, D. Liu, L. Dianli, L. Hui, W. Weizhong, and Z. Kun (Harbin Medical University, Harbin, Heilongjiang 150001, China).

From October 1986 to September 1992, 1,290 cases of birth defects were registered among 160,272 newborns in 25 hospitals in 13 different cities and counties of Heilongjiang province. The overall prevalence was 8.05%. Among all kinds of birth defects were 633 cases of congenital malformations of the central nervous system (255 in males, 363 in females, and 15 for which the sex of the newborn could not be identified). These were the most common birth defect (49.07%), with a prevalence of 3.94%. These 633 cases comprised 496 cases of neural tube defects (78.35%), including anencephaly (249 cases), spina bifida (166 cases), encephalocele (54 cases), and meningocele (27 cases); 130 cases of hydrocephalus; three cases of microcephaly; and four others. The month with highest prevalence rate of congenital malformations of the central nervous system was January. Among the cases of congenital malformations of the central nervous system, were 350 cases of birth weight of <2,500 g (55.29%); 274 cases in which the pregnancy term was less than 31 weeks (43.29%); and 462 cases in which the mother's educational level was less than middle school (72.99%). The data in this study will play an important role in the study of the etiology of congenital malformations of the central nervous system and in developing strategies to prevent these malformations.

Methods I
Chair: Scott Osborne

Selecting an Exposure Lag Period: Is the Model with the Largest Exposure Effect the Best Model? A. Salvan,* L. T. Stayner, and K. Steenland (National Institute for Occupational Safety and Health, Cincinnati, OH 45226).

In epidemiology, there is the inclination to consider more credible the larger estimates of exposure effect. For example, higher relative risks or rate ratios are often emphasized as a criterion for choosing among various hypothesized induction-latency parameters, for example exposure-lag values. Whereas this "higher estimate" approach may often work in practice, the validity of such a criterion is not demonstrated. The purpose of this paper is to use examples

to compare exposure-lag choices based on the highest estimate approach versus those based on a statistical goodness of fit criterion (likelihood ratio test). It seems likely that most epidemiologists believe that the two criteria are equivalent, that is, they would lead to the same choice of exposure-lag parameter estimates, or they may at the most show trivial differences. The examples shown are based on both published and artificial data, in which an exposure-lag parameter is estimated by trial and error fitting: the behavior of the goodness-of-fit statistic obtained over the assigned values of the parameter is compared with that of the relative risk. The examples show that there can be conspicuous inconsistencies between the highest-estimate and likelihood-based goodness-of-fit criteria. In addition, the highest-estimate criterion is shown to lack general validity. It is therefore recommended that, in the absence of a priori biologic knowledge, the selection of exposure-lag values be based on goodness-of-fit criteria.

Effect of Misclassification of Disease Status on Study of Trends. S. Xue,* M. Sprafka, and A. Goldman (University of Minnesota, Minneapolis, MN 55454).

For many diseases, there is no clear diagnostic demarcation. One may choose a set of diagnostic criteria with high sensitivity at expense of specificity or vice versa. Misclassification in either direction always reduces statistical power if the sensitivity and specificity of the diagnostic criteria are the same for each comparison group. To investigate the effects of nondifferential misclassification of disease status on the statistical power of test for trends in binomial proportions, a Monte Carlo study was conducted by simulating various choices one may have between disease classification criteria. According to the study, the loss of statistical power due to misclassification is a function of disease prevalence and ratio of sensitivity and specificity. Diagnostic criteria with equal sensitivity and specificity can minimize power loss only when prevalence is about 50%. For a prevalence lower than 50%, a decline in specificity has a greater impact on statistical power than a decline in sensitivity of equal magnitude. Statistical power loss can be minimized when diagnostic criteria with relatively higher specificity and lower sensitivity are chosen properly. The differential impact of specificity and sensitivity on statistical power becomes considerable when the disease prevalence is lower than 10%. Therefore, it may be inappropriate to adopt diagnoses directly from hospital or clinical records for epidemiologic studies of trends in certain diseases, because of the physician's tendency toward high sensitivity. In contrast, when disease prevalence is higher than 50%, the decline in specificity has less impact on statistical power than a decline in sensitivity. The strategies for applying this technique to choose diagnostic criteria for studies of disease trends are discussed.

Correcting Interclass Correlation Coefficients for Measurement Error in Unbalanced Data Sets. B. Rosner,* A. Wolf, and W. Willett (Harvard University, Boston, MA 02115).

In validation studies, a surrogate measure of exposure (e.g., a food frequency questionnaire) is compared with a gold standard (e.g., dietary records). Often, the gold standard is an average of many replicate measures; using a small number of replicates underestimates the true correlation coefficient. The authors have reported methods to correct the correlation between a single surrogate and multiple replicates of the gold standard using a balanced design (e.g., an equal number of days of diet records for each individual) (*Am J Epidemiol* 1988;127:377–86). Many validation studies, however, have an unbalanced structure, either by design or because of missing data, where a variable number of replicates of the gold standard are available for each subject. A method has been developed to estimate true ρ and associated confidence limits in unbalanced designs, assuming that missing data are missing at random; ρ is estimated using the pairwise estimator (the Pearson correlation based on all pairs $(y_i, x_{i1}) \ldots (y_i, x_{in_i})$ where n_i equals the number of replicates for the ith subject). This estimator has been shown to have minimal bias and high efficiency in previous studies of familial correlation. The authors then correct ρ for measurement error to obtain a deattenuated estimate and associated confidence limits. This procedure is more efficient than correcting only the balanced subset of individuals with complete data. An interactive computer program has been developed to use these methods. These methods are illustrated using replicate daily activity diaries compared with responses to a physical activity questionnaire in the Nurses Health Study II.

Efficient Study Designs for Screening Test Accuracy. L. M. Irwig,* P. P. Glasziou, P. Mock, C. Chock, and J. M. Simpson (University of Sydney, New South Wales 2006, Australia).

Test sensitivity and specificity can be estimated directly on a sample selected independently of test results. If tests are invasive or expensive, the reference (gold) standard, can be applied to all test positives in a screened population and a random sample (the sampled fraction) of test negatives. For this design to result in unbiased estimates and confidence intervals, appropriate corrections must be made for the sampling fraction. Based on the correction formulae of Begg and Greenes (*Biometrics* 1983;39:207–15), the authors have determined the optimal sampling strategy to minimize the total sample size for a given standard error for sensitivity. The scenarios cover prevalences of 1% and 0.1%, sensitivities from 30% to 95%, and specificities from 90% to 99%. Unless sensitivity is very high, obtaining the reference standard for more test positives and undersampling test negatives is often appropriate. The optimal sampling fraction for negatives is lowest when the disease is rare and the test has high specificity. For example, when disease prevalence is 0.1% and test specificity is 99%, the optimal sampling fraction is about 0.3 when the sensitivity is 90%, and 0.1 if the sensitivity is 50%. The savings in overall sample size are 4% in the former case and 39% in the latter. Optimal sampling strategies for sensitivity do not materially increase confidence intervals for specificity.

L-Shaped Dose-Response between Ethanol and Risk of Coronary Heart Disease, from a Refutationist Approach to Meta-Analysis. M. Maclure* (Harvard University, Boston, MA 02115).

More than 20 hypothesized explanations for the

ABSTRACTS OF PAPERS PRESENTED AT THE
TWENTY-SIXTH ANNUAL MEETING

of the
Society for Epidemiologic Research

Keystone, Colorado
June 16–18, 1993

CONTRIBUTED PAPER SESSIONS

Hazardous Exposures and Reproductive Outcomes
Chair: Cheryl Blackmore

Subclinical and Clinical Pregnancy Loss: Incidence and Risk Factors in a Cohort of Working Women. R. B. Hakim,* R. H. Gray, and H. A. Zacur (The Johns Hopkins University, Baltimore, MD 21205).

A cohort of 148 women employed in the semiconductor industry participated in a longitudinal study of reproductive function from 1988 to 1991. Participation involved collection of daily urine samples, completion of daily diaries, and monthly telephone interviews. Daily urine specimens were assayed for steroid hormones to determine menstrual cycle function. Urinary concentrations of human chorionic gonadotropin (hCG) were measured to detect subclinical pregnancy for 675 menstrual cycles using an immunoradiometric assay that detected hCG levels as low 0.05 ng/ml with nearly 100% specificity. These subclinical pregnancies were identified by a rise in hCG of at least 0.25 ng/ml for two consecutive days near the expected time of implantation. Of the 92 pregnancies identified, 34.8% were subclinical losses. The total rate of post-implantation pregnancy loss, including recognized miscarriage, was 50.0%. Factors that adversely influenced pregnancy loss included a history of miscarriage (odds ratio (OR) = 3.2, 95% confidence interval (CI) 1.0–11), nulliparity (OR = 3.5, 95% CI 0.9–14), fertility treatment during the study (OR = 4.5, 95% CI 1.1–27), and smoking (OR = 1.4, 95% CI 0.3–6.1). Factors that had a protective effect on pregnancy loss were consumption of caffeine (OR = 0.4, 95% CI 0.1–0.9) and alcohol (OR = 0.6, 95% CI 0.3–1.3).

Prospective Study of Menstrual Cycle Outcomes in Semiconductor Workers. E. Gold,* B. Eskenazi, M. O. Rasor, K. Hammond, B. Lasley, S. Samuels, and M. Schenker (University of California, Davis, CA 95616).

Previous work has shown that women working in silicon wafer fabrication (FAB) have an increased risk of miscarriage. To determine whether FAB work might also adversely effect menstrual cycles, 156 FAB workers and 252 nonfabrication (non-FAB) workers at seven sites of five semiconductor companies were enrolled for up to six menstrual cycles of follow-up using questionnaires and daily diaries. Exposure assessment was performed by study industrial hygienists. The adjusted mean cycle length (MCL) was 0.64 days longer in FAB than in non-FAB women ($p = 0.38$); women in the thin film/ion implant (TFII) work group had the longest MCL (34 days) as compared with non-FAB women (31.3 days) ($p = 0.06$). The adjusted relative risk for any long cycles (>35 days) was 1.17 (95% confidence interval (CI) 0.81–1.51) and for any short cycles (<24 days) was 1.03 (95% CI 0.70–1.46). The TFII group also had the highest risk for having all cycles be long (adjusted relative risk = 2.59, 95% CI 0.90–6.47). The adjusted geometric mean of standard deviation in cycle length (MSD) was 0.60 days longer in FAB than non-FAB women ($p = 0.12$), with FAB supervisor/engineers having the greatest MSD (4.4 days) as compared with non-FAB women (3.9 days) ($p = 0.17$). While these test results must be interpreted with caution because of multiple comparisons and the limitations of the eligibility criteria, they support other study findings of reduced conception rates in workers in the dopefilm super group.

Occupationally Related Magnetic Field Exposure and Male Subfertility. L. S. Lundsberg,* M. B. Bracken, and K. Belanger (Yale University, New Haven, CT 06510).

Epidemiologic studies of environmental exposures and adverse health outcomes often cite individuals engaged in occupations exposed to electromagnetic fields as being at high risk. A nested case-control study was conducted among nonvasectomized men pursuing infertility diagnosis at the Yale New Haven Hospital Infertility Clinic from 1984 to 1987. Using parameters of morphology, motility, and concentration

---

* Indicates person who delivered the paper or presented the poster.