

001

Maternal Neuroticism and Low Birth Weight in a Nationally Representative US Sample. Mengxiong Wang, Yaqiong Zhu, Jing Jin (University of Florida, Public Health Department, Gainesville United States)

Background: Neuroticism corresponds to the sensitivity of the negative affect system. A person high in neuroticism is defined as someone who is more prone to depression, anxiety consciousness, and demonstrates high emotional reactivity to stress. Some previous studies reported the relationship between maternal personality and pregnancy outcomes, and some pointed out that mothers who were neurotic were more likely to have adverse pregnancy outcomes, such as preterm birth or low birth weight infants. **Objectives:** This study aims to measure the association between maternal neuroticism and low birth weight as a pregnancy outcome, among young women, aged 24-32 living in United States by using the Big Five Inventory. Further, this study also tests whether the level of neuroticism differs by levels of household income. **Methods:** We used the data from the first-wave-interview and fourth-wave-interview of the National Longitudinal Study of Adolescent Health (Add Health) to estimate the logistic regression model. The Big Five Inventory was used to calculate the scores of neuroticism, and the women were classified into either neurotic group or non-neurotic group based on their neuroticism scores. **Results:** There were 4,420 women who participated this research (mean age: 29 years old), and 50.3% of them had yearly household income less than \$50,000. The results showed that the adjusted odds ratio (adjusted for tobacco use, education, race, and sleep pattern) between maternal neuroticism and low birth weight among low income group is 1.39 (95%CI: 1.06-1.84; p-value: 0.02), while among high income group is 0.80 (95%CI: 0.54-1.19; p-value: 0.27). **Conclusion:** Maternal neuroticism is significantly related to low birth weight among low income group. Among women who were in the high income group, the maternal neuroticism was not a significant risk factor. Also, we can conclude that income is a strong modifier in the association between maternal neuroticism and low birth weight.

003

The Why and How of Pooling: Making Models that Make Sense. Emily Mitchell, Robert Lyles, Amita Manatunga, Enrique Schisterman (National Institute of Child Health and Human Development, Bethesda, United States)

Pooling involves the physical combination of biospecimens into a single composite sample prior to performing laboratory assays. Popularized during World War II as a cost-saving tool when testing military recruits for syphilis, pooling has demonstrated additional benefits with the increasing pervasiveness of biospecimen analysis. Pooling can facilitate epidemiological sub-studies by requiring only a portion of each specimen in order to achieve sufficient volume for analysis, thus increasing the number of potential analyses for a fixed number of collected specimens. Despite these benefits, researchers may be hesitant to adopt a pooling strategy since appropriate statistical methods are still being developed, and existing methods for individually-measured specimens may not directly apply to pools. This is particularly true when a biomarker is treated as the outcome in a regression model, since measurements are often positive and right-skewed. Current methods for analyzing this type of data are either computationally expensive or limited to specific pool types. In this study, we propose a novel, flexible and accessible estimation technique for a right-skewed outcome subject to pooling, regardless of pool type. We use simulations to demonstrate the efficacy of our proposed method compared with existing methods. While naive analysis or improper application of existing methods produces estimates with up to 38% bias, our proposed method produces unbiased estimates with confidence interval coverage ranging from 93.8 to 95.2. These simulations, along with analysis of data from the Collaborative Perinatal Project (CPP), demonstrate that when appropriate estimation techniques are applied to strategically-formed pools, valid and efficient estimation can be achieved. This novel method contributes to the base of available statistical tools to analyze pooled specimens and will help empower researchers to more confidently consider pooling as a potential study design.

002

Smoking and Trajectories of Dysmenorrhea among Young Australian Women. Hong Ju, Mark Jones, Gita Mishra (University of Queensland, Brisbane, Australia)

Objective: To investigate the association of cigarette smoking at baseline and trajectories of dysmenorrhea in a large sample of Australian women. **Design** A prospective cohort study. **Setting** Australian (population-based survey). **Participants** A total of 9,067 young women, with at least three measures of dysmenorrhea, randomly sampled from national Medicare database and followed up from 2000 to 2012. **Main outcome measures** Trajectories of dysmenorrhea. **Results:** At baseline, approximately 25% reported dysmenorrhea and 26% were current smokers. Four trajectory groups were identified for dysmenorrhea: normative (42%), late onset (11%), recovering (33%) and chronic (14%), with the chronic group showing high probabilities of reporting dysmenorrhea over time. Compared to never smokers, a significantly higher odds of being in the chronic group was detected for smokers, with odds ratios (ORs) 1.33 (95% confidence interval (CI) 1.05, 1.68) for ex-smokers and 1.41 (95% CI 1.17, 1.70) for current smokers, after adjusting for socio-demographic, lifestyle and reproductive factors. An inverse relationship was identified for earlier age of smoking initiation, with the respective ORs of 1.59 (95% CI 1.18, 2.15), 1.50 (95% CI 1.18, 1.90) and 1.26 (95% CI 1.03, 1.55) for initiation of smoking ≤ 13 , 14-15 or ≥ 16 years. No consistent relationship was evident between smoking behaviour and the odds of being in the other trajectory groups.

004

Truncation Investigations – Using Sensitivity Analyses to Uncover Potential Sources of Bias Due to Truncation in Time to Event Studies. Emily Mitchell, Karen Schliep, Enrique Schisterman (National Institute of Child Development and Human Health, Bethesda, United States)

Truncation is a prevailing issue in many time-to-event studies, but can easily be overlooked since truncated samples, by definition, are never observed. When truncation is not properly accounted for, effect estimates can exhibit bias, particularly when the probability of truncation is differential between exposure groups. In this study, we seek to analyze time to pregnancy (TTP) among women enrolled in the Effects of Aspirin in Gestation and Reproduction (EAGeR) trial. We calculate the fecundability risk ratio (FR) for women who waited more than three months after a pregnancy loss before trying to become pregnant again (Inter-trying interval (ITI) > 3 months), compared with those who waited less than three months (ITI < 3 months), where a FR > 1.0 denotes a shorter TTP. Since eligibility for inclusion into the EAGeR trial was limited to women who were not pregnant at time of enrollment, women with a short TTP would be more likely to be truncated. Differential truncation could occur if women with ITI < 3 months were more likely to have a shorter TTP, and thus were more likely to be excluded from the sample. To investigate the effects of truncation, we compare a naive survival analysis of the EAGeR data to one in which truncation is properly accounted for. While the naive model results in a FR of 1.15 (CI: 0.94 to 1.41) for women with ITI < 3 versus ≥ 3 months, adjustment for truncation results in a FR of 1.63 (CI: 1.32 to 1.99). Based on these results, we conduct a sensitivity analysis to mimic these observed effects of truncation. Sensitivity analyses demonstrate that ignoring truncation produces FR estimates that are biased by 28%, while properly accounting for truncation via delayed entry models effectively eliminates this bias. These results, in conjunction with the data analysis, shed light on possible sources of truncation in the EAGeR trial and emphasize the importance of adjusting for truncation in order to obtain valid inference in time to event studies.—Abstract submitted live 11/5/2014

005

Evaluating epidemiologic influences on cancer incidence trends in a well-defined subpopulation. Hilary A. Robbins, Meredith S. Shiels, Ruth M. Pfeiffer, Eric A. Engels (Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, United States)

Trends over time in cancer incidence rates within a subset of the general population can result from three broad epidemiologic influences: changing 1) changing demographic structure of the subpopulation (e.g., aging), 2) changing general population (background) cancer incidence rates, and 3) changes in the relative risk for cancer associated with membership in the subpopulation. These influences may act independently or in combination to produce the overall incidence trend. We characterized cancer trends during 1996-2010 among HIV-infected people as a subset of the U.S. population. We used Joinpoint regression to divide 1996-2010 into intervals where the slope across annual cancer incidence rates was log-linear. We then evaluated the influence of each factor by 1) comparing regression coefficients for calendar year before and after adjustment for demographic structure, 2) evaluating trends in general population incidence rates after standardization to the HIV population, and 3) evaluating demographics-adjusted trends in the standardized incidence ratio (SIR), which quantifies the HIV-associated relative risk for cancer. We considered each influence to be acting where, respectively, 1) the regression coefficients differed significantly from each other, 2) the trend in standardized background incidence was significant, and 3) the trend in the SIR was significant (p

007

Toward a Better Portrayal of Confounding Bias in Instrumental Variable Applications. John Jackson, Sonja Swanson (Harvard, School of Public Health, Boston, United States)

Background: Recommendations for reporting instrumental variable (IV) analyses often include presenting the balance of covariates across levels of the instrument and levels of the treatment. However, such presentation can be misleading as relatively small imbalances of covariates across levels of the instrument can result in larger bias due to bias amplification. **Methods:** We describe the bias due to omitting a confounder from an IV vs. non-IV analysis, and introduce covariate balance plots that incorporate the strength of the instrument into its scale. Plots are demonstrated using previously published data on proposed preference-based, geography-based, and distance-based instruments. **Results:** Scaled plots provide more accurate context of the bias due to omitting a covariate from an IV vs. non-IV analysis. Scaled plots also provide relevant comparisons of different proposed instruments considered in the same data. **Conclusions:** Covariate balance plots can be adapted to account for bias amplification in the assessment of confounding in instrumental variable analyses.

006

Does it look like a sequentially randomized trial? Covariate balance in studies of time-varying and other joint-exposures. John Jackson (Harvard, School of Public Health, Boston, United States)

Epidemiology often involves analyzing exposures that vary over time, and in recent years methods have been developed for situations when common causes of time-varying exposures and outcomes are affected by prior exposure (i.e. time-dependent confounding). Yet few diagnostics are available to assess the degree of confounding by such time-dependent risk factors in traditional analyses that improperly adjust for such factors, or to demonstrate how well g-methods such as marginal structural models or the longitudinal g-formula address such confounding by emulating a sequentially randomized trial. These complex methods involve several layers of parametric decision-making with consequences not readily apparent to research consumers. We extend approaches for evaluating covariate balance to the setting of time-varying exposures that naturally encompass other joint effects, including mediation and interaction. Balance measures appropriate for longitudinal data are explicitly cast in terms of the assumptions needed to draw causal inference using counterfactual statements and directed acyclic graphs. Intuitive covariate balance plots are then developed to diagnose potential time-dependent confounding in crude data and evaluate how well g-methods balance time-varying risk factors. Using simulated data, we will explore how the conceptual framework and tools developed herein can be used to diagnose time-varying confounding, evaluate the performance of modeling decisions, and detect residual bias in the widely used marginal structural model approach and a recently proposed alternative strategy involving longitudinal g-computation over propensity score strata. This work represents a novel and practical advance in providing tools for transparent reporting of potential confounding in longitudinal data, evaluating the consequences of parametric decisions and post-hoc adjustments, and demystifying the application of sophisticated methods to longitudinal data. —Abstract submitted live 11/5/2014

008

Perception about working in rural area after graduation and associated factors: A study among final year medical students in medical schools in Southeast Nigeria. Uchenna Chidi Anyanwagu, Edmund N Ossai, Ben Azogu, Kenechi Uwakwe, N C Ibiok, Ngozi Ekeke (University of Nottingham, Anambra State Nigeria)

Introduction: Globally, there is an uneven distribution of all cadre of health workers between the urban and rural areas. This is most pronounced in Sub-Saharan Africa with far reaching implications. This study aims to determine the perception of rural medical practice among final year medical students in medical schools in southeast Nigeria and the associated factors. **Method:** A descriptive cross-sectional study was conducted among all final year medical students in Southeast Nigeria, using self-administered questionnaire. Information was obtained on their demography and reasons/willingness to practice in the rural areas after graduation. Data analysis was done using STATA version 13. Bivariate and multivariate logistic regression analysis was done to identify the factors associated with their willingness to practice in the rural area after graduation. **Results:** 457 students participated, representing a response rate of 86.7%. The mean age was 25.5±2.9 years. 13.6% were willing to practice in rural areas after graduation of which 38.7% hoped to serve there for less than 10 years. Majority of the students (80.1%) were of the opinion that doctors in the rural area should earn more than their urban counterparts. Factors associated with willingness to practice in rural areas included family residence in urban area, (OR=0.44, 95% CI: 0.21-0.93); work experience before admission into medical school (OR=2.03, 95% CI: 1.01-3.85); intention to specialize in Community Medicine (OR=3.06, 95% CI: 1.21-7.74) and satisfaction with rural posting (OR=2.14, 95% CI: 1.17-3.90) after adjusting for age, type of institution, educational attainment of mother, sponsorship of university education and either or both parents being a medical practitioner. **Conclusion:** Majority of the students were unwilling to work in rural area after graduation. Knowing the need for doctors here, a re-orientation of students on rural practice through adequate community exposure during rural posting is essential.

009

Exposure to Alcohol Outlets in Rural Towns. Chris Morrison (Monash University, Department of Epidemiology and Preventative Medicine, Melbourne, Australia)

Alcohol related problems tend to occur more frequently in areas with greater concentrations of alcohol outlets, and greater concentrations of outlets tend to be found in lower income areas. This study examines whether the economic processes common to all retail markets might explain this health disparity for residents of rural towns, because outlets are attracted to towns with greater resident and temporary populations due to greater demand but are excluded from higher income towns due to land and structure rents. Using data from the 2011 Census in Victoria, Australia, a population-density based algorithm aggregated the smallest available spatial units into 353 discrete towns (mean population = 4,326.0, SD = 15,754.1). Static demographic data and a matrix of one-day commute journeys provided the independent variables that characterized each town, including the income for the local and adjacent towns, the total resident population, and the net changes to population (due to commuter flow, visitors, and the estimated flow of local residents to other towns). Bayesian conditional autoregressive Poisson models then predicted counts of bars, restaurants and off-premise outlets within the towns. Lower local income and increased income in adjacent towns were associated with more outlets of all types. Greater resident populations and greater net population due to commuters also predicted greater numbers of all outlets. Bars and restaurants were positively related to greater net population due to visitors, and negatively related to the flow of residents away from towns. Cumulatively, these relationships suggest that residents of lower income towns are exposed to excess risk associated with exposure to additional outlets that service demand from residents of other towns.

011

Fixed versus random effects models for longitudinal data analysis of confounding in ecological time series: a simulation study. Usama Bilal, Thomas A. Glass (Johns Hopkins Bloomberg School of Public Health)

Background: Analysis of ecological panel data in the econometrics literature has usually been conducted using fixed effects (FE) models due to their capacity to deal with time-fixed confounding. Random-effects (RE) models are more prevalent in the epidemiologic literature. Group-level centering of covariates can be a useful technique to eliminate time-fixed confounding in RE models. **Objective:** To study if group-level centering of covariates in random-effects models eliminates time-fixed confounding. **Methods:** We conducted a simulation study comparing fixed-effects and random effect models (with and without group-level centering of covariates). We started with a sample of gross domestic product (GDP) from 100 countries for 50 years and randomly selected a subset of 30 countries. A time-fixed binary variable correlated with GDP was created. We then modeled mortality trends in each country with a quadratic underlying trend and the effect of GDP and the binary confounder. Different scenarios with increased confounding were tested. We simulated 1000 datasets for each scenario and computed average percent bias and average mean squared error. **Results:** Under no confounding, all models showed almost no bias (Average Percent Bias: FE: 0%, RE: 7%, Centered RE: 0.17%). Increased levels of confounding resulted in increased differences in the level of bias. In the scenario with the highest levels of confounding, FE and Centered RE showed good performance compared to RE (Averaged Percent Bias FE: 0%, RE: 30%, Centered RE: 0.07%). Both RE models showed much higher accuracy than FE models. **Discussion:** Random-effects models with group-level centering of covariates shares the advantages of fixed-effects models (deals with time-fixed confounding) while keeping the advantages of random-effect models and showing much higher accuracy than fixed-effect models. Studies of ecological time trends in epidemiology should consider the use of this type of models.—Abstract submitted live 11/5/2014

010

Breast size and breast cancer risk: a systematic review. Mengxiong Wang, Chandni Pate, Lusine Yaghjian (University of Florida, College of Public Health Professions and College of Medicine, Gainesville, United States)

Background: Because cell number and division rate can increase cancer risk, it was suggested that breast size can increase breast cancer risk. However, the evidence on the association between breast size and BC is inconsistent. **Objective:** This systematic review aims to summarize published studies that investigated the association between breast size and BC risk, and to explore the underlying causes of variation in the results across the studies. **Methods:** This review included studies published in English between January 1st, 1990 and March 31st, 2014 that were accessible in full-text format. A search was conducted using the terms “breast size (or bra size)” and “breast cancer (or breast neoplasms)” through PubMed Central, BioMed Central, Embase, and Springer One. Manual checks of reference lists from the retrieved articles were used to identify additional manuscripts. **Results:** We found one cohort and seven case-control studies that satisfied our inclusion criteria. Four studies reported a significant association between breast size and BC risk. The findings were compared while stratifying the studies by sample size, women’s menopausal status, and breast size measurement approach. We observed no differences in the associations by these study characteristics. The studies that reported significant findings all had racially diverse study populations, but none of them reported on the association separately by race/ethnicity or adjusted for it. **Conclusions:** There is no strong evidence in support of the association between breast size and breast cancer risk. Significant associations may be limited to specific characteristics, such as BMI

012

Mathematical Modeling for Longitudinal Analysis of Mood Outcomes in Bipolar Disorder; Applying Fourier Series for Handling Non-Linearity. Shervin Assari, Masoud Kamali, Melvin McInnis (University of Michigan, School of Public Health, Ann Arbor, United States)

Bipolar disorder is one of the least studied psychiatric disorders [1-3]. This is particularly due to complexity and non-linearity of change of mood outcomes over time in this disorder. Modeling linear slope may be an oversimplistic approach and may result in model miss-specification. Many studies of longitudinal trajectories of mood outcomes among patients with bipolar disorder ignore time, and consider variance, frequency and maximum of symptoms over a certain time interval. This approach is also of limited accuracy as time of an event is important and should not be ignored. In response to the above listed issues, there is an increasing interest in application of sophisticated mathematical models to mood longitudinal changes of the outcomes in bipolar disorder [4]. There are previous studies that have used nonlinear oscillator model and dynamical systems to describe mood changes in bipolar disorder [5-7]. Although Fourier series have been used for handling non-linearity among patients with bipolar disorder [8-12], very few studies - if any - have ever used this model to explain changes of mood symptoms of bipolar patients. Current study was an attempt to use Fourier series for explaining longitudinal changes in mood outcomes among patients with bipolar disorder. Data came from the Prechter Bipolar Cohort with about 500 participants and up to 5 years of follow up. Participants were diagnosed as having different types of bipolar disorder. Symptom of depression (PHQ) and mania (YRMS) were measured every 2 months. Conventional fit indices were used to evaluate goodness of fit of the model with the data. Our findings suggest that Fourier series may be a useful tool for explaining longitudinal changes of depressive and manic symptoms over time among patients with bipolar disorder. These findings may have public health and clinical implications, as bipolar disorder is one of the most severe psychiatric disorders and influences 4% of general population in the U.S..