

## Poster Abstracts

### **1. Missing Exposure Data in Stereotype Regression Model: Application to Matched Case-Control Study with Disease Subclassification**

Jaeil Ahn ([jaeil@umich.edu](mailto:jaeil@umich.edu)), Bhramar Mukherjee, Stephen B. Gruber, and Samiran Sinha

With advances in modern medicine and clinical diagnosis, case-control data with characterization of finer subtypes of cases are often available. In matched case-control studies, missingness in exposure values often leads to deletion of entire stratum, and thus entails a significant loss in information. When subtypes of cases are treated as categorical outcomes, the data are further stratified and deletion of observations becomes even more expensive in terms of precision of the category-specific odds-ratio parameters, especially using the multinomial logit model. The stereotype regression model for categorical responses lies intermediate between the proportional odds and the multinomial or baseline category logit model. The use of this class of models has been limited as the structure of the model implies certain inferential challenges with non-identifiability and non-linearity in the parameters. We illustrate how to handle missing data in matched case-control studies with finer disease subclassification within the cases under a stereotype regression model. We present both a Monte Carlo based full Bayesian approach as well as an expectation/conditional maximization algorithm for estimation of model parameters in presence of a completely general missingness mechanism. We illustrate our methods by using data from an ongoing matched case-control study of colorectal cancer. Simulation results are presented under various missing data mechanisms and departures from modeling assumptions.

### **2. Multiple Imputation in Group Randomized Trials: The Impact of Misspecifying Clustering in the Imputation Model**

Rebecca R. Andridge ([randridge@cph.osu.edu](mailto:randridge@cph.osu.edu))

In group randomized trials (GRTs), identifiable groups of subjects (e.g. clinics, communities, or schools) rather than individuals are randomized to study conditions, while the outcomes of interest are observed on individuals within each group. Resulting data consist of a small number of groups (clusters), each containing a relatively larger number of subjects, and within each group observations are likely to be correlated. Usually the resulting intracluster correlation (ICC) is small, with typical values in the 0.05 to 0.001 range, but even small values can lead to large variance inflation factors and cannot be ignored. Missing data is also a problem with GRTs, and multiple imputation is often used to create complete data sets that can be analyzed using well-established analysis methods for GRTs. I discuss strategies for accounting for clustering in multiple imputation for a missing continuous outcome, focusing on a simple post-test only design where the intervention effect is tested with an adjusted two-sample t-test. The two-sample t-test requires imputation with a mixed effects model (exchangeable correlation within a cluster) in order for the analysis to be congenial to the imputation model; however this type of imputation is not yet available in standard statistical software packages. An alternative approach that could be carried out in standard software is to include fixed effects for cluster, but the impact of this misspecification has not been studied. I show that under this type of imputation model the multiple imputation variance estimator is biased and, somewhat counter-intuitively, smaller ICCs lead to larger

overestimation of the multiple imputation variance. Analytical expressions for the bias of the variance estimator are derived in the case of missing completely at random (MCAR), and the case in which data are missing at random (MAR) is illustrated through simulation. Finally, various imputation methods are applied to data from the Detroit Middle School Asthma Project, a recent school-based GRT, and differences in inference are compared.

### **3. Inference on direct and mediated causal effects using the sequential regression multiple imputation framework**

Irina Bondarenko ([ibond@umich.edu](mailto:ibond@umich.edu)) and Trivellore Raghunathan

Causal Inference, formulated using potential outcomes, can be viewed as a missing data problem. The same framework can also be used to partition the population averaged causal effect into This work presents a way to approach causal analysis and partitioning total effects into the direct and mediated effects. Our goal is to illustrate the use of sequential regression multiple imputation to estimate the direct and mediated causal effects using the locally developed software, IVEWARE. We adopt the concept of principal stratification and treated unobserved potential outcome as missing data and multiply impute them. Monotonicity constraints are imposed in the imputation process. We compare the estimates from this approach to those based on fully Bayesian approach implemented using MCMC. We consider cases where the treatment is binary and mediator and outcomes are any combinations of binary/continuous/count/semi-continuous variables. We incorporate baseline covariates through propensity score stratification. The method is illustrated using NHANES data.

### **4. Variable Selection for Multiply-Imputed Data**

Qixuan Chen ([qc2138@columbia.edu](mailto:qc2138@columbia.edu)) and Sijian Wang

Multiple imputation is frequently used to handle the missing data. However, there are currently no guidelines for the extension of commonly used variable selection methods for complete data to the setting of multiply imputed data. We proposed two variable selection methods for multiply-imputed data, which lead to a single selected model across the different imputed data sets. The first method is a modification of the traditional stepwise variable selection method, implemented by obtaining combined p-values using Rubin's multiple imputation combining rule first and then selecting variables base on combined p-values in each step of selection. The second method is based on the regularized joint-likelihood estimation approach. The coefficients of the same variable across all imputed data are treated as a group, and the group lasso penalty (Yuan and Lin 2006) is applied for the purpose of variable selection. Due to the penalty term, all models from corresponding imputed datasets will be fitted *jointly*. Furthermore, owing to the nature of group lasso penalty, for some variables, *all* of the estimated coefficients of the same variable will be *exactly* zero. Simulation studies show that for important variables with large regression coefficients, both the modified stepwise method and the group lasso method can identify such variables as important with more than 95% of time. For important variables with small regression coefficients, the group lasso method performs much better than the modified stepwise selection. However, the group lasso method has larger positive false rate, that is, it is more likely to identify noise variables as important.

## **5. A method to estimate the sharing of eQTLs between tissues, with application to skin and lymphoblastoid cells**

Jun Ding ([junding@umich.edu](mailto:junding@umich.edu)), Johann E. Gudjonsson, Liming Liang, Philip E. Stuart, Wei Chen, Rajan P. Nair, James T. Elder and Gonçalo R. Abecasis

Gene transcript levels can serve as an intermediate phenotype that bridges genotypes and more complex organismal phenotypes, including common diseases. Genome-wide association studies of gene expression in several human tissues have identified thousands of genetic loci impacting the expression of specific transcripts. Each of these loci is called an expression quantitative trait locus (eQTL). Although it is expected that many eQTL will be tissue specific, the exact proportion of eQTL that are tissue specific or shared between tissues remains unknown.

A simple measure of the tissue specificity of eQTLs can be obtained by examining the overlap of eQTL lists from two different tissues. Unfortunately, this naive approach will always underestimate the true proportion of overlapping signals. We have developed a more accurate method. Our multi-step procedure first generates a list of potential eQTLs and then uses unbiased estimates for QTL effect sizes to estimate the expected number of replicating eQTLs for a specific sample size. The proportion of overlapping eQTL can then be interpreted in this context. When applied to compare cis-eQTL detected in an analysis of 57 skin biopsies and in a panel of ~400 lymphoblastoid cell lines, our method shows that 70-80% of eQTL are shared between tissues, a much larger proportion than the naive estimate of 30-40%.

Our results provide guidance to researchers contrasting eQTL results across tissues and a specific means to accurately estimate the proportion of overlapping eQTL between tissues.

## **6. A Markov Compliance Behavior and Outcome Model for Causal Analysis in the Longitudinal Study**

Xin Gao ([xingao@umich.edu](mailto:xingao@umich.edu)) and Michael R. Elliott

We propose a Markov compliance behavior and outcome model for analyzing longitudinal randomized studies when non-compliance is present. We solve the problem in the potential outcome framework, and provide causal estimates on the treatment effect via principal stratification. Previous research considered the effect of subjects' joint compliance behavior on the joint distribution of the longitudinal outcomes, but we consider the effect of compliance behavior at time  $t-1$  on the potential outcomes at time  $t$ , and the effect of treatment effect at time  $t$  on the compliance behavior on time  $t+1$ . The proposed model provides estimates of the ITT effect within each principal stratum, and the effect of the treatment effect on the following adherence. We analyze the Suicide CBT Study using our proposed model via Markov chain Monte Carlo (MCMC) methodology.

## **7. Regression Analysis on Covariates that Have Heteroscedastic Measurement error**

Ying Guo ([guoy@umich.edu](mailto:guoy@umich.edu)) and Roderick Little

We consider the problem of estimating the regression of an outcome  $D$  on a covariate  $X$ , where  $X$  is unobserved, but a variable  $Y$  that measures  $X$  with error is observed. A calibration sample that measures pairs of values of  $X$  and  $Y$  is also available; we consider the internal and external calibration sample based on including values of  $D$  or not. One common approach for measurement error correction is Regression Calibration (RC), substituting the unknown values of  $X$  by predictions from the calibration curve of  $X$  on  $Y$ . An alternative approach is to multiply impute the missing values of  $X$  given  $Y$  and  $D$  based on an imputation model, and then use multiple imputation (MI) combining rules for inferences. A recent paper by Freedman et al (2008) compares these two approaches, suggesting that RC is more efficient under plausible assumptions. However, their work assumes the measurement error of  $Y$  has a constant variance, whereas in many situations, the variance varies as a function of  $X$ . We consider modifications of the RC method and the MI method that allow for heteroscedastic measurement error, and compare them by simulation. The MI method is shown to provide better inferences in this setting.

## **8. Joint Modeling of Survival and Binary Endpoints via Semiparametric Transformation Models**

Chen Hu ([chenhu@umich.edu](mailto:chenhu@umich.edu)) and Alexander Tsodikov

Survival and binary endpoints which are jointly observed and correlated require special attentions in modeling. For example, stage-specific cancer incidence represents a random vector of joint bivariate response represented by the age at diagnosis, and cancer stage. How to link the observed stage-specific cancer incidence with unobserved tumor progression and history of metastasis before diagnosis are of particular interests. In this paper, the joint response is modeled through a series of semiparametric models with time-dependent covariates. We extend the framework of semiparametric transformation model inference proposed by Chen, Jin, and Ying (Biometrika 2002), and modified partial likelihood approach proposed by Bagdonavicius and Nikulin (Life Data Analysis 1999) to establish estimation and inference procedures. This method is illustrated by simulation studies and prostate cancer data from the Surveillance, Epidemiology and End Results (SEER) program.

## **9. Modeling Menstrual Cycle Lengths at the Approach of Menopause Using Bayesian Change Point Models**

Xiaobi Huang ([xiaobih@umich.edu](mailto:xiaobih@umich.edu)) and Michael Elliott

As women approach menopause, the patterns of their menstruation segment lengths change. In order to study when changes in menstrual length happen, we build Bayesian linear change point model to model the mean segment length as well as the variability of the length. The full model includes separate linear change points for each woman and a hierarchical model to link them together, along with regression components to include predictors of menopausal onset such as age at menarche and parity. Data are from TREMIN, an ongoing 70-year old longitudinal study that has obtained menstrual calendar data of women throughout their life course. Our study cohort include nearly 1000 women, many of

whom have missingness due to hormone use, surgery, random missingness and loss of contact. We integrate multiple imputation in our Bayesian estimation procedure to deal with different forms of the missingness. Bayesian model checking has been performed to assess the fit of the model.

## **10. Meta Analysis of Functional Neuroimaging Data via Bayesian Spatial Point Processes**

Jian Kang ([jiangkang@umich.edu](mailto:jiangkang@umich.edu)), Timothy D. Johnson, Thomas E. Nichols and Tor D. Wager

There is a growing interest in meta analysis of functional neuroimaging studies. Typical neuroimaging meta analysis data consist of peak activation coordinates (PACs) from several studies. Most published methods only produce null-hypothesis inferences and do not provide interpretable, fitted model. To overcome these limitations, we propose a Bayesian hierarchical marked spatial Cox cluster process model. The posterior intensity function provides information on the most likely locations of population centers as well as the inter-study variability of PACs about the population centers. We model the PACs as offspring of latent realizations of a study center process for each study. And the study-level point processes are the offspring of latent realizations of a population center process. To reduce the bias in results, our model incorporates weights for each study, based on the quality of the study, as marks of the process. We illustrate our model with a meta analysis consisting of 437 studies from 164 publications and study our model via sensitivity analyses and simulation studies.

## **11. Treatment Effects under Early Detection of Cancer**

Shih-Yuan (Connie) Lee ([shihylee@umich.edu](mailto:shihylee@umich.edu)) and Alex Tsodikov

Cancer-specific survival counted out from the point of diagnosis is the most common endpoint used in cancer clinical trials and observational studies. Estimated treatment effect in proportional hazards models may be biased if models are misspecified. We consider a misspecification associated with ignoring the random mechanism of the early detection of the disease. We study how the lead time resulting from the early detection of cancer affects the treatment effect as estimated by a proportional hazards model ignoring the advanced diagnosis. We study how the multiplicative treatment effect differs in screened vs. non-screened population and assess the magnitude and the direction of the bias under various conditions. We also investigate the behavior of the estimator in the Cox proportional hazards model using partial likelihood. A meta-analytic approach is proposed to correct the bias and variance of Cox PH estimators based on a joint cancer incidence and survival modeling approach.

## **12. A Permutation Test for Random Effects in Linear Mixed Models**

Oliver Lee ([oel@umich.edu](mailto:oel@umich.edu)) and Tom Braun

Inference regarding the inclusion or exclusion of random effects in linear mixed models is challenging because the variance components are located on the boundary of its parameter space under the typical null hypothesis. As a result, the asymptotic null distribution of the Wald, score, and likelihood ratio tests will not have a chi-squared distribution under the null hypothesis. When the null distribution includes  $p$  random effects and alternative hypothesis contains  $p+1$  random effects, it has been proved that the appropriate asymptotic distribution is a 50/50 mixture of two chi-squared distributions with  $p$  and  $p+1$

degrees of freedom. However, this approach suffers from two limitations. First, this reference distribution is known to lead to conservative tests, i.e. those with size below their nominal level, in small samples. Second, the appropriate mixture distribution is rather cumbersome and non-intuitive when the null and alternative hypotheses differ by more than one random effect. As an alternative, we present a permutation test whose statistic is a sum of weighted squared residuals, with the weights determined by the among- and within-subject variance components. The null permutation distribution of our statistic is computed by permuting the residuals both within- and among-subjects and is valid both asymptotically as well as in small samples. We examine the size of our test via simulation in a variety of settings and compare it to the size based upon the classical mixture-of-chi-squares approach.

### **13. Causal Assessment of Surrogacy in a Meta Analysis of Colorectal Clinical Trials**

Yun Li ([yunlisph@umich.edu](mailto:yunlisph@umich.edu)), Jeremy Taylor, Mike Elliott, and Bhramar Mukherjee

When the true endpoints (T) are difficult or costly to measure, surrogate markers (S) are often collected in clinical trials to help predict the treatment effect (Z). There is a vast interest in understanding the relationship among S, T and Z. Traditional models have been used to assess surrogacy; however, these models often condition on a post-randomization variable S which may cause bias. A counterfactual-based principal stratification concept has been proposed by Frangakis and Rubin (2000) to study their causal associations. In this paper, we propose a Bayesian estimation method to assess surrogacy in a multiple trial setting and obtain the trial-specific and overall causal associations among S, T and Z in this counterfactual framework. The method allows information sharing across trials and improves the estimation precision. All S, T and Z are binary. We extend our method to the settings with and without the monotonicity assumption. We assess the method using simulations. We then apply it to two colon cancer examples and evaluate the goodness-of-fit of the models. Our research shows the results are very sensitive to the monotonicity assumptions.

### **14. Genome-Wide eQTL Analysis: Cross-Tissue Comparison of Brain and Lymphoblastoid Cell Lines**

Ellen Schmidt ([schellen@umich.edu](mailto:schellen@umich.edu)) and Margit Burmeister

Poor availability of the primary tissue of interest when studying human brain disorders has led to the widespread use of more easily obtainable sources such as lymphoblastoid cell lines (LCLs). While this tissue source is invaluable for many studies, questions remain as to what extent gene expression regulation differs between brain and LCLs. We investigate this using a genome-wide expression quantitative trait loci (eQTL) mapping technique in both human brain and lymphoblastoid cell lines. As part of an ongoing collaborative effort with the Pritzker Neuropsychiatric Disease Consortium, postmortem brain samples from a study cohort of approximately 100 unrelated individuals are analyzed. Microarrays are used to quantify mRNA expression levels from 6 different regions of the brain, taken from controls or cases with one of 3 psychiatric disorders including bipolar disorder, schizophrenia, and depression. SNP data from these samples are generated using a 550K Illumina chip, with an additional 2 million SNPs imputed based on known genotypes from the International HapMap Project. The study is comprehensive, analyzing approximately 2.5 million imputed DNA polymorphisms, or SNPs, against expression levels of 22,000 transcripts. LCL data are obtained from publicly available results of the study

on 60 CEU unrelated HapMap individuals, and contain SNP information for the same 2.5 million SNPs and expression data for 46,000 transcripts. We present and compare preliminary results of the association computations for 6 different brain regions and lymphoblastoid cell lines.

### **15. Improving Small-Sample Inference in Group Randomized Trials with Binary Outcomes**

Philip Westgate ([pwestgat@umich.edu](mailto:pwestgat@umich.edu)) and Tom Braun

Group Randomized Trials (GRTs) randomize groups/clusters of people to treatment or control arms instead of individually randomizing subjects. Typically, GRTs have a small number of independent clusters,  $n$ , each of which can be quite large. As a result, any inferential method whose properties appeal to asymptotics is not appropriate in this setting. We deal with scenarios in which each subject has a binary outcome, i.e. “success” or “failure”, and randomizing groups or clusters of subjects creates the potential for over-dispersed binomial data. The over-dispersion is created by the variability between-groups, i.e. subjects within the same cluster may have a different probability of a success as compared to subjects from other groups, even if their respective groups were assigned to the same study arm. The degree of variability can be quantified by the intra-cluster correlation coefficient (ICC). Assuming the ICC is a nuisance parameter, analyses can be done using quasi-likelihood with a logistic link. A Wald statistic, which asymptotically has a standard normal distribution, can be used to test for a marginal treatment effect. However, this asymptotic result may not be valid as the sample size  $n$  can be quite small in most GRTs. We deal with the fact that in small samples, a Wald test may have a test size that is smaller than its nominal value because the statistic has a variance less than one. When the ICC is known, we develop a method for adjusting the estimated standard error appropriately such that the Wald statistic will approximately have a standard normal distribution even when  $n$  is small. We also propose ways to handle non-nominal test sizes when the ICC is estimated. Asymptotically, our proposed methods are equivalent to using the ordinary Wald statistic with standard normal critical values. Through simulation results covering a variety of realistic settings for GRTs, we examine the small-sample performance of our methods.

### **16. Predicting Treatment Efficacy via Quantitative MRI: A Bayesian Joint Model**

Jincao Wu ([jincaowu@umich.edu](mailto:jincaowu@umich.edu)) and Timothy D. Johnson

Despite success in the treatment of many cancers, the prognosis for patients with high-grade gliomas is poor, with a median survival of one-year post diagnosis. Furthermore, the assessment of therapy efficacy in these patients does not occur until approximately 5-6 months after the commencement of therapy. Salvage therapies, if warranted, are not prescribed until after this assessment has taken place and are typically too late to have any substantial impact on overall survival. Our colleagues hypothesize that quantitative MRI (qMRI) can reduce this time, to assess treatment efficacy, to three weeks after therapy begins, thereby allowing salvage treatments to begin earlier when first-line therapies fail. The purpose of this work is to build an early predictive model of treatment efficacy using qMRI data and assess its performance. We use one-year survival status as the outcome and propose a joint, two-stage Bayesian model. In the first stage, we smooth the qMRI data using a spatio-temporal pairwise-difference prior (PWDP). We propose four, novel summary statistics that are calculated from the smoothed images

in the first stage. In the second stage, these statistics enter into a generalized non-linear model (GNLM) as predictors of survival status. In our GNLM we use the probit link and a Multivariate Adaptive Regression Spline (MARS) basis. Gibbs sampling and reversible jump Markov chain Monte Carlo are applied iteratively between the two stages to estimate the posterior distribution. Cross validation is employed to assess the performance of our model in predicting one-year survival. Through both simulation studies and model performance comparisons we find that we are able to attain lower overall misclassification rates by accounting for the spatio-temporal correlation in the images and by allowing for a more complex and flexible decision boundary (than that afforded by a standard probit regression model) provided by the GNLM.

### **17. Psychological Correlates of Placebo Responsivity in Healthy Controls**

Hamdan A. Yousuf ([yousufh@umich.edu](mailto:yousufh@umich.edu)), Christian S. Stohler and Jon-Kar Zubieta

**Objective:** To examine the degree to which placebo response is associated with psychological and behavioral characteristics, including state and trait aspects.

**Methods:** Thirty-seven healthy young volunteers (20 males, 17 females, mean age:  $25.57 \pm 4.23$ ) completed a set of standard personality questionnaires and underwent positron emission tomography scans with the  $\mu$ -opioid-receptor-selective radiotracer carfentanil labeled with carbon 11. Measures of receptor concentrations were obtained while the subjects underwent a 20-minute standardized pain challenge, in the absence and presence of a placebo with expected analgesic properties. Placebo response was defined as the percent difference in pain rating between the two conditions.

**Statistical Analyses:** Correlations were computed between placebo response and selected variables of *a priori* interest from among the psychological questionnaires. Ordinary least squares regression models were fit for placebo response and the most significant univariate predictors were identified. Due to multicollinearity in the predictor variables, a partial least squares regression model using leave-one-out cross validation was chosen to simultaneously optimize predictive power for the response and explanatory power for the variance in the predictor space.

**Results:** While expectations of placebo effectiveness were not a significant predictor of response, difference between expectations and subjective rating of efficacy obtained post-trial were highly significant. Ego-resiliency and agreeableness were the most powerful univariate positive predictors of response, while neuroticism was a negative predictor. The model weights for the PLS regression reveal the importance of trait well-being in general. The overall variance explained by the model is nearly 30%, which provides significant evidence of trait and state effects on placebo responsivity.

### **18. Methods for rare risk variants in the age of re-sequencing**

Matthew Zawistowski ([mattz@umich.edu](mailto:mattz@umich.edu)) and Sebastian Zöllner

Results from genome wide association studies (GWAS) illustrate that common genetic variants are insufficient to explain the heritability of most common traits. Therefore, it is hypothesized that rare variants are major contributors to common complex diseases. Next generation sequencing methods

have greatly improved the ability to discover and type rare variants and have made resequencing of large samples an affordable option. However, current common variant methods that test individual markers for an association with the phenotype of interest have insufficient statistical power to identify rare risk variants. Here, we introduce a novel method, the Combined Minor Allele Test (CMAT), for identifying rare risk variants by testing entire genes for an excess of rare variants among cases. The CMAT identifies, within the sequencing data, a set of candidate sites defined as rare polymorphic sites predicted to be deleterious by bioinformatics tools. The CMAT then compares the cumulative number of minor alleles at the candidate sites across the entire gene between cases and controls. A statistically significant CMAT result indicates a gene in which cases harbor more copies of rare, deleterious variants than do controls.

In order to compare the performance of CMAT to existing rare variant methods, we simulated case-control sequence data under a model of allelic heterogeneity. We find that across numerous study scenarios the CMAT proves to be the most powerful test of those considered. Notably, under non-ideal study conditions such as low sample size and high variant misspecification, the power gain for CMAT over other methods is most significant.

Another advantage of our method is that it can be applied if only a subset of the data is sequenced and the rest is genotyped for a dense marker panel, e.g. if candidate genes identified in a GWAS were subsequently sequenced in a subset of individuals from that study. By imputation of those rare variants that were detected by resequencing, we can infer most genotypes in the entire sample and use those genotypes in our test statistic. We optimize an existing imputation method (MACH) for rare variants and estimate imputation error rates. Based on these estimated error rates we illustrate the power gained by including imputed genotypes in CMAT. Finally, we demonstrate how our method can be applied to existing GWAS data augmented by imputing from HAPMAP and present results from a re-analyzed GWAS.