

## 2009 Biostatistics Student Research Day Abstracts

Monday August 10, 2009 (Alumni House)

### **Cost and accuracy comparisons in medical testing using sequential testing strategies**

**Anwar Ahmed**

**Faculty advisors: D. McClish/C. Schubert**

The practice of sequential testing is followed by the evaluation of accuracy, but often not by the evaluation of cost. We compared three sequential testing strategies: believe the negative (BN), believe the positive (BP) and believe the extreme (BE), which is a new strategy introduced in this research as an alternative to the BP and BN strategies. All three strategies were used to combine two medical tests to diagnose a disease or condition. Comparison between these strategies will be performed in terms of accuracy and cost, in order to minimize the number of tests needed for each subject. As such, we offer solutions for limiting healthcare costs by introducing a new graphical technique incorporating cost called the Minimum Cost Maximum Receiver Operating Characteristic (MCMROC) curve. Within these strategies, four different parameters were identified that could influence the performance of the combined tests: the area under curve (AUC) of each individual test, the ratio of disease and non-disease standard deviations ( $b$ ), independence ( $\rho$ ) between disease positive negative results, and prevalence. Simulation was used to compare cost and accuracy between the strategies. An example assuming  $\rho=0$  and  $b=1$  will be discussed. The majority of situations show that one of the strategies always has better performance. Rare situations show that two of the strategies give the same results, so the choice of which strategy to use will depend on the context of the test. We found that the cost could be reduced further by setting an acceptable tolerance on test sensitivity.

### **Using patient data as a QC tool: Does a reagent change affect similar shifts in patient and QC assays**

**Tina Cunningham**

**Faculty advisor: Robert E. Johnson**

In reporting laboratory assay outcomes, an important goal is to ensure that results for patients are consistent across different reagent lots. Good laboratory practice requires each new lot of reagent is verified before being put into service. It is well documented that results from quality control (QC) materials are subject to bias caused by different matrix interactions with different reagent lots. A preferred approach is to examine patient results that have been measured using correctly calibrated old and new lots of reagent.

This project investigates the suitability of QC results to verify consistent results for patient samples when changing reagent lots. Laboratory QC and patient data records from 11/15/2005 to 02/03/2009 were provided, including 5 analyzers, 83 analytes, 468 reagent lot changes, the number of patients and QC concentrations used in each reagent lot change evaluation. Ten QC measures were analyzed prior to and 10 were analyzed post each change. Results for 5-10 patient samples were measured prior to and post each change.

The computational method calls for some deviations from the traditional t-test. The pre-post QC samples are independent with assumed equal variance. Paired data for each patient were available, thus a difference was computed.

This computed difference has an assumed equal variance across patients but different from the QC variance. An extension from the Satterthwaite approximation was used to compare the difference between QC and patient measurements. Initial results showed that 34.8% of the QC measurements are significantly different from the patient sample measurements.

### **Identification of determining factors on childhood obesity**

**Amber Wilk**

**Faculty advisor: Yongyun Shin**

Obesity among children between six and eleven years of age has not only has accelerated to an epidemic rate in recent years, but is a key predictor of obesity in adulthood. In this logistic analysis, I attempt to identify the determinants of child obesity and to shed light on their policy implications. Children's socioeconomic status (SES), both individually and environmentally, has been found to be negatively associated with the obesity. Because children, unlike adults, spend a majority of their time at school, their obesity may potentially be hugely impacted by their school environmental factors. This analysis controls for the compositional effect of the school characteristic, school mean SES, on obesity, along with individual characteristics such as racial/ethnic differences and age, and finds major determinants of a child's body mass index that may have important policy implications in order to reduce obesity among children.

### **Statistical methods for methylation beadarrays**

**Maria Capparuccini**

**Faculty advisor: Kellie J. Archer**

DNA methylation is an epigenetic modification that has been implicated in various cancers. It is defined as the addition of a methyl group to the 5'-carbon of a cytosine located next to a guanine and is typically restricted to CpG islands. The Illumina GoldenGate Methylation BeadArray Cancer Panel I platform is technology developed for cancer-focused methylation analysis. This is a two-channel platform whereby a DNA sample undergoes bisulfite conversion followed by extension and ligation using primers for the methylated and unmethylated target sequence, and methylation status is then commonly measured as the proportion of methylation measured at each CpG site. An interrogated CpG site is then determined to be differentially methylated in a cancerous tissue by means of a two sample t-test with a null hypothesis of equality in proportion methylated between the cancerous and non-cancerous tissue. While currently a widely accepted method for determining differentially methylated CpG sites, application of the two sample t-test to proportion data presents a clear violation of the distributional assumptions which motivate this test. Therefore, an alternative method is explored based on a bivariate gamma distribution of the methylated and unmethylated signal intensities which underlie the proportion methylated variable.

Preliminary results will be described for a dataset consisting of 36 liver samples. Among these 36, sixteen of the samples presented hepatitis C virus (HCV) cirrhosis without concomitant hepatocellular carcinoma (HCC) and 20 presented HCV-cirrhosis with concomitant HCC. Interest was in identifying differentially methylated CpG sites comparing HCV cirrhosis with and without concomitant HCC.

### **Detection call algorithms for high-throughput gene expression microarray data**

**Sarah Reese**

**Faculty advisor: Kellie J. Archer**

Prior to analyzing microarray data, various summarization methods are applied to the raw signal intensities for obtaining gene-level summaries, often in a platform-specific manner. For Affymetrix GeneChip data, the detection call algorithm is a summary method based on the Wilcoxon signed-rank test that declares whether a given transcript is present or absent in a given sample, and is commonly used to assess the reliability of expression levels (Pepper et al, 2007) as well as to filter out transcripts that are not truly differentially expressed (Hackstadt and Hess, 2009). For the Illumina gene expression platform on the other hand, although a detection call method has been implemented in their BeadStudio software, it has not been shown that it can reliably assess expression levels and filter transcripts, nor have specific thresholds been identified for designating a transcript as present or absent. In this talk, an Illumina spike-in data set will be described and the Illumina BeadStudio detection call method will be presented. In addition, an alternative detection call method for Illumina data, developed based on a standard t-test method, will be described and compared to the current Illumina method.

## **An inferential approach for one-sample gene network comparisons: assessing dissimilarity via a resampled local covering metric**

**Phillip Yates**

**Faculty advisor: Nitai Mukhopadhyay**

The analysis of weighted co-expression gene sets is gaining momentum in bioinformatics circles. In addition to research directed towards inferring co-expression networks inferential methods are being developed to compare transcription networks across one or more phenotypes. Common gene set hypothesis testing procedures are mostly confined to comparing the average gene/node transcription levels between one or more groups and make limited use of additional network features. Ignoring the gene set architecture disregards relevant network topological comparisons and can result in familiar  $n < p$  overparameterized test issues.

Here we propose to compare a sampled co-expression network with a target network where the measure of separation is determined with a local covering metric. The proposed metric is an additive measure over each of the genes/nodes in the network. Comparable to other local dependency structures, e.g., spatial or AR(1) models, the dissimilarity measure at each node is determined using the network properties of nearby neighbors. Since the cover primarily relies on edges, weights, and in- and out-degrees the specification of a more complex network parameterization is avoided. In order to draw statistical inferences we use a resampling approach. Our method, which admittedly discounts large-distance or clustered co-expression effects, allows for both an overall network test and an examination of individual gene/node effects. In addition to testing for network relational differences our dissimilarity statistic can be extended to incorporate the mean/variance transcription comparisons of existing gene set methods. We evaluate our proposed metric using both simulated data and a widely-analyzed diabetes microarray data set.

## **Effects of chemical mixture exposure on hormone networks**

**Stephanie Pearson**

**Faculty advisor: Roy Sabo**

**Background:** Due to the involvement in many human biological processes, exposure to endocrine disrupting chemicals (ECDs) can pose numerous health risks. While the general relationships between some of these types of chemicals and hormonal behavior have been established, the internal bodily mechanisms behind these relationships are much less studied.

**Objective:** It is of interest to analyze the relationships between several types of polychlorinated EDCs (bi-phenyls (PCBS), di-benzo dioxins (dioxins) and di-benzo furans (furans)) and hormones (follicle stimulating hormone (FSH), luteinizing hormone (LH), thyroid stimulating hormone (TSH) and thyroxine (T4)), while accounting for certain internal

bodily characteristics (such as plasma glucose, cholesterol and triglycerides). Data is utilized from participants of the 2001-2002 National Health and Nutrition Examination Survey (NHANES), a cross-sectional survey of a representative sample of the U.S. population.

**Methods:** Multi-level modeling is used to assess the effects of internal bodily characteristics on the relationships between the specified EDCs and both the reproductive (FSH and LH) and metabolic (TSH and T4) hormones. Covariates include factors such as menopausal status, age, gender, and others. Hormone concentrations are dichotomized into normal and abnormal in order to bypass inherent skewness.

## **The influence of race/ethnicity on neurobehavioral problems in individuals with traumatic brain injury**

**Caroline Carr**

**Faculty advisor: Jessie Ketchum**

**Objective:** To determine differences in NFI (Neurobehavioral Functioning Inventory) subscales (depression, somatic pain, memory, communication, aggression, and motor functioning) at 1-year post traumatic brain injury (TBI) among White, Black, and Hispanic individuals, after adjusting for demographic and injury characteristics.

**Methods:** 1,237 individuals from the TBI Model Systems National Database with moderate to severe TBI (895 Whites, 272 Blacks, and 70 Hispanics) occurring between 1996 and 2001. ANCOVA models were fit for each subscale and included an effect for race/ethnicity, as well as covariates that differ between the races/ethnicities, significantly affect outcome, or modify the effect of race/ethnicity on outcome. Higher NFI scores represent more problems.

**Results:** The NFI scores were different among the races/ethnicities for all six subscales; however the differences were modified by covariates. Hispanics with non-violent injuries had greater depression, somatic, and motor scores than Blacks or Whites with non-violent injuries. Whites with violent injuries had greater depression scores than Hispanics with violent injuries. Hispanics who were young at injury (20 years) had greater memory scores than Blacks who were young. Whites employed at injury had greater memory scores than Hispanics employed at injury. Blacks and Hispanics who were unmarried had greater communication scores than Whites who were unmarried. Whites with high disability at discharge (10) had greater aggression scores than Blacks with high disability. Blacks with low disability at discharge (2.5) had greater motor scores than Whites with low disability.

**Conclusions:** Neurobehavioral treatment programs for Hispanics and Black individuals with TBI need to be designed and implemented.

## **Investigating Euclidean distance as a similarity measures for detecting sufficient similarity in dose response in clinical methods: A child care center study.**

**Scott Marshall**

**Faculty advisor: Chris Gennings**

Using statistical equivalence testing logic and mixed model theory an approach has been developed, that extends the work of Stork et al (JABES,2008), to define sufficient similarity in dose-response for chemical mixtures containing the same chemicals with different ratios or a subset of chemicals. The proposed similarity measures are based on Euclidean distance. The method is applied to a child care center study (Tulve et al., 2006) investigating exposure to pesticides. The example demonstrates the ability of the proposed method to be used for the purpose of environmental risk assessment and further establishes surrogate measures to evaluate risk. *(Partially supported by NIEHS #T32 ES007334 and does not reflect USEPA policy. This research is not associated with Monsanto.)*

**Tuesday August 11, 2009 (Almuni House)**

## **Evaluation of the International/Inner City / Rural Preceptorship (I<sup>2</sup>CRP) program**

**Yan Jin**

**Faculty advisor: Robert E. Johnson**

**Objective:** This survey study will evaluate the influence on medical students' empathy and understanding of patient medical beliefs and attitudes after enrolling in VCU's International / Inner City / Rural Preceptorship program (I<sup>2</sup>CRP).

**Methods:** The Jefferson Scale of Physician Empathy (JSPE) and the Health Beliefs and Attitudes Survey (HBAS) will be applied for evaluation at three time points during the study period. The study population will include the pre-doctoral students of VCU's School of Medicine who enroll in the I<sup>2</sup>CRP in each of the three study years (N≈60, 20 per year). Students who do not enroll will serve as a comparison group.

**Results:** Summary statistics will be generated at each time point for both JSPE and HBAS scores. The study hypothesis is that the participant group and comparison group will have comparative survey scores at baseline, but the participant group will have higher scores compared to comparison group at follow-up and for the participant group itself, follow-up score will be higher than baseline's score. The relationship of JSPE and HBAS will be performed at baseline overall and separately for two groups at follow-up. Only the baseline results are presented here.

**Discussion:** I will discuss the experience I gained with survey data from a statistician's point view. This includes how to develop a statistical method from the major objectives of the study and apply it. As always there are many steps needed to prepare the data for analysis and these will be outlined.

## **Comprehensive analysis of progressively downregulated genes in liver cancer**

**Tobias Guennel**

**Faculty advisor: Kellie J. Archer**

**Motivation:** Liver cancer is estimated to be the 6<sup>th</sup> leading cancer type and the 3<sup>rd</sup> leading cause of cancer death worldwide. Hepatocellular carcinoma (HCC) is a primary malignancy of the liver and most cases of HCCs are secondary to hepatitis-B virus (HBV), hepatitis-C virus (HCV), or hepatic cirrhosis. Gene expression profiles from 116 independent liver samples using the Affymetrix GeneChip technology, including 19 normal, 57 HCV+ cirrhotic non-HCC, and 40 HCV+ cirrhotic HCC tissues were available for analysis.

**Purpose and Methods:** The goal of this project was to utilize a combination of statistical and bioinformatic methods to identify progressively downregulated genes between the above liver tissues and subsequently analyze their corresponding promoter regions. Specifically, a linear model was fit at a probe set level to the gene expression data and contrasts were used to identify significantly downregulated probe sets while controlling the FDR at 10%. Promoter regions for all genes on the Affymetrix U133a2 gene chip with an entry in Database of Transcriptional Start Sites (DBTSS, [www.dbss.com](http://www.dbss.com)) were extracted and examined for CpG islands.

**Results:** Among the 21091 probe sets tested, 157 probe sets, corresponding to 141 genes, were found to be progressively downregulated. DBTSS contained confident promoter regions for 114 of these genes of which 78.1% contained at least one CpG island within the corresponding promoter region compared to 68.7% of the remaining genes (p=0.02).

**Conclusion:** The findings of our comprehensive analysis of genome-wide gene expression profiles and methylation in a HCC population warrant further methylation studies in HCC.

## **Logit-based and constraint-based models**

**Andre Williams**

**Faculty advisor: Kellie J. Archer**

In most biostatistics-related studies, there is an outcome variable to be modeled. In statistical modeling, the goal is to predict or gain a deeper understanding of the relationship between an outcome variable and a set of independent variables measured on the observational units. In the case of an ordinal outcome, models based on the multinomial distribution are commonly used. These models are also known as logit-based models as they model log odds, or pseudo log odds, as a linear combination of the independent variables. Examples of models commonly used to predict an ordinal response include the baseline category logit, adjacent category logit, Continuation ratio logit, Proportional odds logit, and Stereotype logit models. As powerful as these techniques are, two major limitations present themselves; these procedures cannot model data where there are more variables than observations and these procedures can produce unstable estimates when the parameter estimates have a complex covariance structure. A possible solution is to modify the above models by adding some constraint or penalty. Two common constraint-based models are the Ridge Regression model and the Least Absolute Shrinkage and Selection Operator (LASSO). These constraint-based methods have been applied in the linear, logistic, and survival modeling settings and have demonstrated a variety of benefits in the analysis of high-throughput genomic data. In this talk, an overview of ordinal modeling methods and the LASSO constraint-based method will be provided.

## **Bayesian approaches for estimation of tolerable region in multiple endpoints and multiple hazards exposure**

**Epiphanie Nyirabahizi**

**Faculty advisor: Chris Gennings**

Determining Benchmark Dosages (BMD) is of interest to toxicologists and risk assessors. Often data come from experiments with multiple chemicals and multiple endpoints. Current methodology evaluates each chemical and endpoint separately resulting in multiple statistical tests consequently inflating Type I error rates. Methods to adjust for multiple comparisons such as Bonferroni correction are subject to reducing the power to detect effects of interest. We introduced a Bayesian approach for both multiple endpoints and multiple chemical data into a single unified model. This model is estimated using MCMC in WinBugs and R. Inferences on the endpoint model and chemical effects are done via Bayes' factors. Using this method, and the benchmark dose method, a 95% estimation of a Bayesian' safe exposure region is computed.

## **Deriving optimal composite scores: relating observational/longitudinal data with a primary endpoint**

**Rhonda Ellis**

**Faculty advisor: Chris Gennings**

In numerous clinical/experimental studies, multiple endpoints are measured on each subject. It is often not clear which of these endpoints should be designated as of primary importance. The desirability function approach is a way of combining multiple responses into a single unitless composite score. The response variables may include multiple types of data: binary, ordinal, count, interval data. Each response variable is transformed to a 0 to 1 unitless scale with zero representing a completely undesirable response and one representing the ideal value. In desirability function methodology, weights on individual components can be incorporated to allow different levels of importance to be assigned to different outcomes. The assignment of the weight values are subjective and based on individual or group expert opinion. In this dissertation, it is our goal to find the weights or response variable transformations that optimize an external empirical objective criterion. In particular, we find the optimal weights/transformations that minimize the generalized variance of a prediction regression model relating the score and response of an external variable in pre-clinical

and clinical data. For application of the weighting/transformation scheme, initial weighting or transformation values must be obtained then calculation of the corresponding value of the composite score follows. Based on the selected empirical model for the analyses, parameter estimates are found using the usual iterative algorithms (e.g., Gauss Newton). A direct search algorithm (e.g., the Nelder-Mead simplex algorithm) is then used for the minimization of a given objective criterion i.e. generalized variance.

## **Comparisons of normalization methods on four major microarray**

**Jiayi Hou**

**Faculty advisor: Mark Reimers**

Normalization is an important procedure in microarray analysis. The purpose of normalization is to remove the systematic technical differences between chips in order to see clearly the systematic biological differences between samples.

We use four different normalization methods on samples from the MicroArray Quality Control (MAQC) project, using the four different major microarray platforms: Affymetrix; Agilent; Illumina and Nimblegen. We studied four normalization methods: global median normalization; simple loess normalization; quantile normalization and semi-linear normalization methods. We assessed the effectiveness of the normalization methods using the known relations among the samples; specifically we used an F-test comparing the variations between samples versus the variations among sample replicates.

## **Psychological impact of class II relationships in the mixed dentition**

**Adam Sima**

**Faculty advisor: Al Best**

**Purpose:** The purpose of this study was to determine if a Class II dental and skeletal appearance is perceived by teachers, mothers and peers to predict a child's athletic, social, leadership and academic abilities, and a child's susceptibility to teasing.

**Methods:** A survey was constructed using frontal nonsmiling, frontal smiling and profile nonsmiling photographs of patients (subjects) between ages 8-10. Five subjects had Class II malocclusions, overjet greater than 6mm, lip incompetency and retrognathic profiles (retrognathic), and the other five had Class I skeletal relationships, minimal overjet and lip competency (normal). Teachers, mothers, and peers (N = 368 evaluators) were asked to evaluate the subjects based upon 5 characteristics: athletic ability, leadership ability, popularity, academic ability and susceptibility to teasing. The ratings were recorded digitally by sliding a bar along a 100mm Visual Analog Scale (VAS). A multivariate mixed model accounted for the dependencies of the ratings made by multiple evaluators across the five characteristics. A Bonferroni correction was used to compare differences in evaluator and subject characteristics within each category.

**Results:** Results revealed that retrognathic subjects are perceived to have significantly lower athletic ability and popularity and higher susceptibility to teasing as compared to subjects with a normal jaw ( $P's < 0.001$ ).

**Conclusions:** Mothers, teachers, and peers perceive children with Class II malocclusions, overjet greater than 6mm, lip incompetency and retrognathic profiles are perceived more negatively with respect to athletic, popularity, and teasing than children with Class I skeletal relationships, minimal overjet and lip competency.

**The onset of maturation marks changes in body composition and metabolic risk factors for boys and girls**  
**Bob Carrico**  
**Faculty advisor: Shumei Sun**

The onset of puberty may influence body composition and related cardiovascular and metabolic risk factors differentially in boys and girls. Serial records of self-reported Tanner stages of puberty were examined in 133 boys and 83 girls participating in the Fels Longitudinal Study. Probit analysis was conducted on each child's Tanner stage data in order to model the progression of puberty and estimate the onset of the pubertal growth spurt (PGS). Model fit was assessed overall and compared to sex-specific national estimates for the mean ages of the PGS. Boys and girls were categorized separately as early or late maturers if their estimated ages at the onset of the PGS were below or above the 50th percentile of the sex-specific group estimates, respectively.

Then, using the probit models' estimates for the PGS, mixed models were fit separately to serial data of blood pressure, anthropometric measurements, body composition and fasting plasma lipid levels in the same children over a period from 3 years prior to and 3 years after the onset of the PGS. Test slices were made for each year before and after the PGS to compare the early and later maturers grouped according to sex. Adjustments to the alpha level were made for multiple comparisons.

We found that early and late maturers of both sexes had significant differences in body composition and in serial measurements of blood pressure and fasting plasma lipid levels. Progression in puberty is marked by changes in these important risk factors.