Thomas G. Stewart

tgs@email.unc.edu • www.bios.unc.edu/∼tgs

Portfolio of Collaborative Projects

This document provides a summary of my contributions as a statistical consultant to a number of of collaborative research projects.

Post Approval Drug Monitoring

For the past two years, I have worked with a team of medical doctors, data abstraction and monitoring specialists, and other statisticians in an observational, longitudinal, and multi-center study of Hepatitis C therapies known as HCV-TARGET. The purpose of HCV-TARGET is to provide insight about clinical outcomes in a real-world, non-clinical trial population of Hepatitis C patients. This research is valuable because (a) real-world patient populations differ from clinical trial patient populations and (b) treatment practices of physicians may differ from standard clinical trial protocols. Because the FDA recently approved a series of new therapies for Hepatitis C and because of the number of collaborating sites, HCV-TARGET has been in a unique position to provide monitoring and analysis of patient outcomes in a timely fashion. As part of the research team, I have contributed to several oral and poster presentations along with manuscripts in peer reviewed medical journals. (See CV.) My work for these research papers spanned a wide range of tasks: (a) generating usable datasets from the large registry of patient data, (b) defining and executing the data analysis plans for the various research questions, and (c) generating publication ready tables, figures, and study reports. The analyses in the research papers includes time-to-event survival methods, generalized linear models, missing data methods, and causal inference methods for comparative effectiveness results. My time with HCV-TARGET has been a very positive and very involved collaboration with exposure to most every aspect of an observational, clinical study.

Birth Defects and Air Pollution

I have worked with researchers at the EPA in an exploratory study of birth defects and a mother's exposure to air pollution during the first trimester of gestation. This study utilized a number of large datasets: the EPA's Fused Discrete Air Quality Surfaces Data, Texas vital records for births, and the Texas Birth Defects Registry. The air quality data are daily estimates of ozone and fine particulate matter for a grid of cells (12 km $\times$ 12 km) covering all of Texas (for about 7000 cells including adjacent regions). Because the study period was 5 years, the air quality dataset was $\sim$12.8 million observations ($5 \times 365 \times 7000$). The task of combining this exposure data to the case and control data from vital records was the reason I was brought on the collaborative team. Once the data was merged, I performed the exploratory data analysis and interpreted the results in the proper context of an exploratory data analysis. As a result of this collaboration, I plan to pursue research which helps investigators outside of statistics to understand the important and proper role of exploratory analyses, but also to understand the limitations of exploratory analyses. We as statisticians must provide guidance relevant to the increasingly popular big data environment.

Gestational Diabetes and HBA1C

My work with Berggren, Fellow in Obstetrics/Gynecology, is another example of

collaborative research with observational data in a clinical setting. The study evaluates blood biomarker HBA1C as a predictor of poor birth outcomes for pregnant mothers with gestational diabetes. Because HBA1C is not measured in the standard course of care, the mothers measured for HBA1C represented a biased sample. As such, our analysis required appropriate adjustment for selection bias. One result of this study was the reevaluation of traditional biomarkers used for prediction of poor birth outcomes. While further research is needed for confirmation, this study suggested the possibility that traditional biomarker-cut-points separating high-risk patients from low-risk patients should be changed. I am interested in collaborative research which reevaluates standard practice of care with current patient populations. As this study demonstrates, standards of care several decades old may not be suited to today's patients. Because of modern health records, health researchers have the potential to provide data-driven recommendations like never before.

Injury Prevention

Working with NC Public Health, I created a simple user interface and a number of SAS scripts to help public health experts synthesize injury data. The researchers use the interface to create publication-ready reports with a few simple mouse clicks (no SAS programming on the user's side). You can see an example of the output here (link). I am interested in effective statistical reporting and finding ways for collaborators to explore large datasets.