

**“When Disease Mapping Meets Data Privacy”**

**Harrison Quick, Ph.D.**

**Assistant Professor**

**Dept. of Epidemiology and Biostatistics**

**Drexel University**

# Abstract

When agencies release public-use data, they must be cognizant of the potential risk of disclosure associated with making their data publicly available. This issue is particularly pertinent in disease mapping, where small counts pose both inferential challenges and potential disclosure risks. While the disease mapping and statistical disclosure limitation literatures are individually robust, this talk aims to begin to build a bridge between these areas of research.

First, I’ll discuss the notion of generating and disseminating synthetic data for public use.  I’ll then demonstrate how models that account for spatiotemporal dependence structures can be used generate synthetic data for public use using ten years of county-level heart disease death counts for multiple age-groups.  While this approach produces synthetic data with high utility, it is unclear whether it provides formal privacy guarantees – e.g., differential privacy.  To this end, I’ll introduce the concept of differential privacy and provide an overview of my recent work related to generating differentially private synthetic death data in the context of disease mapping.

# 12:00 Noon, Tuesday, April 21, 2020

# 47 College Street, 106B

# 11:45 AM Lunch served outside Rm. 106B

