

---

# STATISTICS SEMINAR

---

UW-Department of Statistics

[www.stat.wisc.edu](http://www.stat.wisc.edu)



**Abstract:** Integrating multiple, complex sources of data is a growing challenge in statistics and machine learning. For example, in the multi-block (or multi-view) data setting we are given two or more data matrices with a fixed set of observations (e.g. patients) and multiple sets of features (e.g. clinical variables and gene expression data). In this setting, Angle-based Joint and Individual Variation Explained (AJIVE) extracts a joint signal common to all data blocks as well as individual signals which are specific to each data block. We give an overview of the core statistical inference procedure underlying AJIVE. We then explore an application to the Carolina Breast Cancer Study which involves H&E stained tumor biopsy images as well as gene expression data. In this application, images are represented as data vectors using Convolutional Neural Networks (CNNs). While the use of neural networks allows us to capture rich visual information, it leads to major problems for interpretability. We present ongoing work to develop techniques to interpret visual modes of variation captured by CNNs. Finally, we discuss new directions for statistical data integration.

**TITLE:** Joint  
Analysis of H&E  
Stained Images and  
Genetic Covariates  
Using Convolutional  
Neural Networks and  
AJIVE

**Speaker:**

**Iain  
Carmichael**

Graduate Student  
University of North  
Carolina Chapel Hill

**Time & Place:**

Monday, January 28,  
2019 **4pm**, Room  
133 SMI  
Cookies & Coffee @  
**3:30**, Rm 1210 MSC

