Statistical Methods in Imaging Conference

June 5-7, 2018

University of Pennsylvania Perelman School of Medicine

June 5, 2018

R Hack-A-Pack Event

8:30a-9:00a	**Breakfast** (provided)	**BRB 14th Floor Lobby**
8:45a-9:00a	**Introduction to hackathon, presentation of the space/schedule**	**BRB 14th Floor Lobby**
9:00a-9:45a	**Lightning talks & tutorials**	**BRB 14th Floor Lobby**
9:45a-10:00a	**Coffee break**	**BRB 14th Floor Lobby**
10:00a-10:30a	**Project pitches**	**BRB 14th Floor Lobby**
10:30a-10:45a	**Team organization**	**BRB 14th Floor Lobby**
10:45a-12:45p	**Open hacking**	**BRB 253 / 1412**
12:45p-1:30p	**Bag Lunch** (provided)	**BRB 14th Floor Lobby**
1:30p-3:00p	**Open hacking**	**BRB 253 / 1412**
3:00p-3:20p	**Project updates**	**Gaulton Auditorium**
3:20p-5:30p	**Open hacking**	**BRB 253 / 1412**
5:30p-7:00p	**Briefings & beers**	**BRB Ground Floor Lobby**

June 6, 2018
Gaulton Auditorium and Lobby
BRB II/III

8:30a-9:00a **Breakfast** (provided in the lobby)

9:00a-10:00a **Overview of Imaging Statistics in R**
John Muschelli, PhD (Johns Hopkins University)

10:00a-11:00a **Brain Connectivity and Parcellation**
Organizer: Amanda Mejia, PhD

Bayesian spatial binary regression for label fusion in structural neuroimaging
Andrew Brown, PhD (Clemson University)

Template ICA: Identifying Brain Networks in Individual Subjects using Empirical Big Data Priors
Amanda Mejia, PhD *(Indiana University)*

Likelihood Based Dynamic Connectivity Analysis using Hidden Semi-Markov Models
Heather Shappell, PhD *(Johns Hopkins University)*

11:00a–11:30a **Coffee Break**

11:30a-12:30p **Statistical Methods for Clinical Imaging: Three Case Studies**
Organizer: Ciprian Crainiceanu, PhD

Consideration on Causal Inference in 4D Flow MRI for Bicuspid Aortic Valve Patients
Adin-Cristian Andrei, PhD (Northwestern University)

Radiomics and imaging of the lung and breast
Nichole Carlson, PhD *(University of Colorado)*

Dynamic prediction of MS lesions: a case for joint functional and survival modeling of voxel trajectories
Ciprian Crainiceanu, PhD *(Johns Hopkins University)*

12:30p-1:45p **Lunch** (provided)

1:45p-2:45p **Multimodal Imaging and Reduction Techniques**
Organizer: Dana Tudorascu, PhD

Multimodal Prediction of Beta Amyloid Load from MRI Brain Images Using Structured Sparse Regression
Joanne Beer, MS (University of Pittsburgh)

Global PCA of Local Moments with Applications to MRI Segmentation
Jake Maronge, MS (University of Wisconsin)

An Integrative Model for Assessing Multimodal Neuroimaging Signatures of Posttraumatic Stress Disorder
Zoe Zhang, PhD (Drexel University)

2:45p-3:15p **Hackathon Report**

3:15p-3:45p **Coffee Break**

3:45p-4:45p **Collaborative Case Study: Background Parenchymal Enhancement in Breast MRI**
Organizer: John Kornak, PhD

Significance of Breast MRI Background Parenchymal Enhancement for Predicting Response to Chemotherapy
Vignesh A Arasu, MD (University of California, San Francisco)

Statistical analysis of MRI of the Breast in the Presence of Background Parenchymal Enhancement
John Kornak, PhD (University of California, San Francisco)

5:00p-7:00p **Poster Reception**

June 7, 2018
Gaulton Auditorium and Lobby
BRB II/III

8:30a-9:00a **Breakfast** (provided in the lobby)

9:00a-10:00a **Analysis and Processing of Complex-Valued MRI**
Organizer: Benjamin Risk, PhD

Statistical impacts of reconstruction method in simultaneous multislice acquisition of MRI
Benjamin Risk, PhD (Emory University)

Bayesian image analysis in Fourier space for Medical Imaging
John Kornak, PhD (University of California, San Francisco)

Bayesian Spatial Modeling via Kernel Convolutions on Complex-Valued fMRI Signals
Cheng-Han Yu, PhD (University of California, Santa Cruz)

10:00a-11:00a **Student Awards Presentations**

11:00a-11:30a Coffee Break

11:30a-12:30p **Collaborative Case Study: Quantitative Immunohistochemistry Biomarkers based on Tissue Microarray Images**
Organizer*: Inna Chervoneva*, PhD

Quantitative immunohistochemistry biomarkers for precision oncology
Hallgeir Rui, MD, PhD (Medical College of Wisconsin)

Spatial statistics approach to develop novel protein cancer biomarkers
Inna Chervoneva, PhD (Thomas Jefferson University)

12:30p-2:05p **Lunch** (provided)

2:05-2:45p **Recent Advances in Modeling Large-Scale Imaging Data**
Organizer: Zoe Zhang, PhD

Statistical modeling of brain networks using multimodal neuroimaging analysis
Ying Guo, PhD (Emory University)

NPBayes-fMRI: Nonparametric Bayesian General Linear Models for Single- and Multi-Subject fMRI Data
Jeong Hwan Kook (Rice University)

2:45-2:55 **Brief Break**

2:55p-3:35p (session continued)

Efficient semi-parametric regression for longitudinal data with regularized estimation of error covariance function
Chunming Zhang, PhD (University of Wisconsin-Madison)

Sparse Multivariate Mediation and Moderated Mediation Analysis
Seonjoo Lee, PhD (Columbia University)

3:35p-4:00p **Panel Discussion**
Haochang Shou, University of Pennsylvania
Dana Tudorascu, University of Pittsburgh
Ted Satterthwaite, University of Pennsylvania
Ipek Oguz, University of Pennsylvania

Abstracts

June 6, 2018

9:00 a.m. – 10:00 a.m.

Overview of Imaging Statistics in R
John Muschelli, PhD (Johns Hopkins University)

Neuroconductor (https://neuroconductor.org) is an open-source platform for rapid testing and dissemination of reproducible computational imaging software. The goals of the project are to: 1) provide a centralized repository of R software dedicated to image analysis, 2) disseminate software updates quickly, 3) train a large, diverse community of scientists using detailed tutorials and short courses, 4) increase software quality via automatic and manual quality controls, and 5) promote reproducibility of image data analysis. We provide a description of the purpose of Neuroconductor, highlight packages in this framework, and some imaging analysis examples with real data sets.

10:00 a.m. – 11:00 a.m.
Brain Connectivity and Parcellation
Organizer: Amanda Mejia, PhD

Bayesian spatial binary regression for label fusion in structural neuroimaging
Andrew Brown, PhD (Clemson University)

Most analyses of neuroimaging data involve studying one or more regions of interest (ROIs) in a brain image. In order to do so, each ROI must first be identified. Since every brain is unique, the location, size, and shape of each ROI varies across subjects. Thus, each ROI in a brain image must either be manually segmented or (semi-) automatically delineated, a task referred to as segmentation. Automatic segmentation often involves mapping a previously manually segmented image to a new brain image and propagating the labels to obtain an estimate of where each ROI is located in the new image. A more recent approach to this problem is to propagate labels from multiple manually segmented atlases and combine the results using a process known as label fusion. To date, most label fusion algorithms either employ voting procedures or impose prior structure and subsequently find the maximum a posteriori (MAP) estimator (i.e., the posterior mode) through optimization. We propose using a fully Bayesian spatial regression model for label fusion that facilitates direct incorporation of covariate information while making accessible the entire posterior distribution. We discuss the implementation of our model via Markov chain Monte Carlo and illustrate the procedure through both simulation and application to segmentation of the hippocampus, an anatomical structure known to be associated with Alzheimer's disease.

Template ICA: Identifying Brain Networks in Individual Subjects using Empirical Big Data Priors
Amanda Mejia, PhD (Indiana University)

Independent component analysis (ICA) is commonly applied to fMRI data to identify resting-state brain networks (RSNs), regions that activate together in the absence of a particular task. Due to high noise levels in fMRI, group-level RSNs are typically estimated by combining data from many subjects in a group ICA (GICA). Subject-level RSNs are then estimated by relating GICA results to subject- level fMRI data. The resulting subject-level estimates are not model-based and often have low reliability. Recently, model-based methods that estimate subject-level and group RSNs simultaneously have been shown to result in more reliable subject-level RSNs. However, this approach is computationally demanding and

inappropriate for small group or single-subject studies. To address these issues, we propose a model-based approach to estimate RSNs in a single subject using empirical population priors based on large fMRI datasets. We develop an expectation-maximization (EM) algorithm to obtain posterior means and variances of subject-level RSNs. We apply the proposed methods to data from the Human Connectome Project and find that the resulting subject-level RSN estimates are significantly more reliable than those obtained from traditional methods.

Likelihood Based Dynamic Connectivity Analysis using Hidden Semi-Markov Models
Heather Shappell, PhD (Johns Hopkins University)

The study of functional brain networks has grown tremendously over the past decade. Most functional connectivity (FC) analyses assume that FC networks are stationary across time. However, there is interest in studying changes in FC over time. Hidden Markov models (HMMs) are a useful modeling approach for FC. However, a severe limitation is that HMMs assume the sojourn time (number of consecutive time points in a state) is geometrically distributed. This may encourage state switches too often. We propose a hidden semi-Markov model (HSMM) approach for inferring functional brain networks from fMRI data, which explicitly models the sojourn distribution. Specifically, we propose using HSMMs to find each subject's most probable series of network states, the cumulative time in each state, and the networks associated with each state. We demonstrate our approach on fMRI data from an anxiety-inducing experiment, where the algorithm was agnostic to alignment and yet discovered the alignment near perfectly. We also present dynamic FC results from Human Connectome Project data where we identify several states associated with different degrees of connectedness within and between sensory, motor, and higher order cognition regions.

11:30 a.m. – 12:30 p.m.
Statistical Methods for Clinical Imaging: Three Case Studies
Organizer: Ciprian Crainiceanu, PhD

Consideration on Causal Inference in 4D Flow MRI for Bicuspid Aortic Valve Patients
Adin-Cristian Andrei, PhD (Northwestern University)

Bicuspid aortic valve (BAV) is a heart disease that affects approximately 1.3% of the population worldwide, with a male predominance of approximately 3:1. The BAV nomenclature refers to the fact that the aortic valve has two leaflets instead of being trileaflet. Common complications of BAV disease are aortic valve stenosis and/or regurgitation, which might eventually require aortic valve replacement surgery. From an imaging standpoint, BAV is characterized by a degradation of the medial elastin fibers of the aortic wall. Wall shear stress (WSS) of the aortic wall is a feature that shows a positive association with the degree of medial elastin fiber degeneration. 4D flow MRI is a powerful noninvasive imaging technique involving 3D phase-contrast MRI over time, which permits the quantification of WSS during the cardiac cycle. Existing 4D flow MRI studies comparing BAV surgical patients to healthy controls have focused on characterizing regional aortic WSS patterns in associative manner. We present potential ways to define a causal effect and provide insights into its practical usefulness.

Radiomics and imaging of the lung and breast
Nichole Carlson, PhD (University of Colorado)

Co-Authors: Fingerlin TE, Fountain K, Ghosh D, Hughes J, Maier L, Russell P, Ryan, S, Warsavage T, Wolverton D, Xing F

The Department of Biostatistics and Informatics at the University of Colorado has recently developed many partnerships to develop novel quantitative phenotypes of the lung and breast. For conditions such as lung and breast cancer along with most lung diseases, the primary diagnostic tool is a CT or MRI image. To date, most of these images are simply

visually scored by expert radiologists. Research shows visual scoring has poor inter-rater reliability and assessment is quite time consuming limiting its use in large scale research. Radiomics and other texture-based analyses have emerged as an approach for quantitatively phenotyping cancer tumors. In one of several examples of Colorado's imaging research, we show how radiomics may be useful for non-invasively identifying the genetic signature of breast cancer tumors. When then show how radiomic approaches easily scale from regions of interest, such as tumors, to quantifying the whole lung. We show that whole lung radiomics are able differentiate different disease patterns in sarcoidosis, an interstitial lung disease. Through the talk we will cover the additional statistical challenges faced in these problems compared to things we have learned from statistical imaging research in the brain.

Dynamic prediction of MS lesions: a case for joint functional and survival modeling of voxel trajectories
Ciprian Crainiceanu, PhD (Johns Hopkins University)

Co-Author: Jordan Johns

We investigate statistical modeling approaches designed for the dynamic prediction of voxel intensities in an MS lesion using longitudinal clinical images. At every time point when a new MRI scan is acquired we are interested in two problems: 1) predict at any time point into the future the intensity of every lesion voxel; and 2) predict whether or not each voxel will be repaired at a particular time into the future. This type of problem naturally leads to the joint modeling of complex, spatially correlated image intensities and survival of the voxel in the "not-yet-repaired" category. We propose a new class of models designed to conduct such modeling and realistic enough to allow for finite-time computational approaches.

1:45 p.m. – 2:45 p.m.
Multimodal Imaging and Reduction Techniques
Organizer: Dana Tudorascu, PhD

Multimodal Prediction of Beta Amyloid Load from MRI Brain Images Using Structured Sparse Regression
Joanne Beer, MS (University of Pittsburgh)

Brain amyloid pathology is an early feature of Alzheimers disease that often develops prior to any cognitive or functional impairment. Brain amyloid can be measured in vivo using the positron emission tomography (PET) tracer Pittsburgh Compound-B (PiB). However, this method is expensive and involves exposure to radiation. Magnetic resonance imaging (MRI), on the other hand, is less invasive. Since different MRI modalities provide different information about brain structure and function, we hypothesized that combining information from various modalities would lead to more accurate assessment of brain amyloid pathology than any one modality alone. In this study, we trained a model using a structured sparse penalized regression estimator, the fused sparse group lasso, to predict total brain amyloid load from a combination of MRI-based imaging modalities. We compare prediction performance using individual modalities alone with the performance of multimodal prediction, and discuss the most predictive features, including specific brain regions and imaging modalities.

Global PCA of Local Moments with Applications to MRI Segmentation
Jake Maronge, MS (University of Wisconsin)

We are interested in describing the information contained in local neighborhoods, and higher moments of local neighborhoods, of complex multimodal imaging techniques at the population level. This is problematic because of the size of medical imaging data. We propose a simple, computationally-efficient approach for representing the variation in

multimodal images using the spatial information contained in all local neighborhoods across multiple subjects. This method achieves 3 goals:
1) decomposes the observed variability images at the population level; 2) describes and quantifies the main directions of variation; 3) uses these directions of variation to improve segmentation and studies of association with health outcomes. To achieve this, we efficiently decompose the observed variation in local neighborhood moments. In order to assess the quality of this method we show results using the 2015 Ischemic Stroke Lesion Segmentation (ISLES) Challenge.

An Integrative Model for Assessing Multimodal Neuroimaging Signatures of Posttraumatic Stress Disorder
Zoe Zhang, PhD (Drexel University)

Post-traumatic stress disorder (PTSD) is a chronic and disabling anxiety disorder that can develop after a person is exposed to a traumatic event. Human neuroimaging provides exciting opportunities to examine structural and functional brain changes specific to PTSD. The use of multimodal neuroimaging is a promising and recent approach to study complex brain disorders by utilizing complementary physical and physiological sensitivities. At the same time, however, the advent of multimodal neuroimaging has brought the need to analyze and integrate neuroimaging data with advanced statistical methods that can make full usage of their informational complexity. Using data from the Philadelphia Neurodevelopmental Cohort (PNC) study, we identify three distinct groups, people with trauma exposure and no PTSD symptoms, people with trauma exposure and long-lasting PTSD symptoms as well as healthy controls. A large number of imaging features from different modalities including MRI, DTI, and resting-state fMRI are derived. We then develop an integrative probabilistic model to combine heterogeneous data from multiple modalities and select predictive multimodal imaging signatures of PTSD.

3:45 p.m. – 4:45 p.m.
Collaborative Case Study: Background Parenchymal Enhancement in Breast MRI
Organizer: John Kornak, PhD

Significance of Breast MRI Background Parenchymal Enhancement for Predicting Response to Chemotherapy
Vignesh A Arasu, MD (University of California, San Francisco)

Co-Authors: Paul Kim, Wen Li, David C Newitt, Ella Jones, Laura J Esserman, Bonnie N Joe, and Nola M Hylton

Background parenchymal enhancement (BPE) represents physiologic uptake of contrast in normal fibroglandular tissue during contrast-enhanced breast MRI. Clinically, BPE is described using four ordinal categories of increased levels of enhancement. It is used to qualify the level of potential masking for diagnosing suspicious lesions on MRI. Moreover, BPE demonstrates both physiologic variability and inter-rater variability. This variability may hamper efforts to use MRI to assess in vivo change in tumor volume when assessing neoadjuvant response. At the same time, recent studies have demonstrated an association between high levels of BPE and increased breast cancer risk. This talk will explore the biological basis of BPE, the clinical assessment using BI-RADS, and efforts in quantitative measurement of BPE. This will provide a the context and motivation for understanding how BPE may confound measurement of tumor volume, but also how BPE may be used as a novel biomarker for predicting response to chemotherapy in an effort to personalize care.

Statistical analysis of MRI of the Breast in the Presence of Background Parenchymal Enhancement
John Kornak, PhD (University of California, San Francisco)

Co-Authors: Karl Young, Nola M Hylton, Bonnie N Joe, Vignesh A Arasu, Ella Jones, David Newitt, Wen Li

Contrast-enhanced structural MRI (breast MRI) is often used for the diagnosis of breast cancer in high-risk patients. Although breast MRI can provide high-resolution imaging with good contrast-to-noise ratio there are a number of difficulties in using these images in clinical studies and also for optimally detecting cancer lesions. Firstly, these images are subject to background enhancement: a complex noise pattern within which patterns can present the illusion of tumors. Secondly, breast MRI varies over time as the contrast injection travels through the body and so the full-time course of images needs to be accounted for. Thirdly, the background enhancement may itself be related to the risk of breast cancer, confounding results. In this talk we will show-case different aspects of analyzing studies of breast MRI breast data. We will look at 1) Bayesian image analysis in Fourier space for modeling the background enhancement; 2) Functional tumor volume measurement models for predicting breast cancer recurrence via percent enhancement (PE) and signal enhancement ratio (SER) approaches; 3) Models that use background enhancement levels as a predictor. This is all work that is under development and we hope to demonstrate that the field is rich in opportunities for new modeling ideas.

June 7, 2018

9:00 a.m. – 10:00 a.m.
Analysis and Processing of Complex-Valued MRI
Organizer: Benjamin Risk, PhD

Statistical impacts of reconstruction method in simultaneous multislice acquisition of MRI
Benjamin Risk, PhD (Emory University)

Simultaneous multislice (SMS) acquisition can be used to decrease the time between acquisition of fMRI volumes, which can increase sensitivity by facilitating the removal of higher-frequency artifacts and boosting effective sample size. The technique requires an additional processing step in which the slices are separated, or unaliased, to recover the whole brain volume. Full field of view images are reconstructed in k-space from multiband slices using kernels estimated with complex-valued least squares (slice-GRAPPA) or constrained complex-valued least squares (split slice-GRAPPA), where the constraints reduce signal leaking into aliased locations. SMS can lead to noise amplification, which decreases the benefits. We conducted a simulation study in k-space to examine the impacts of SMS and reconstruction method. We also examined data from the Human Connectome Project and found evidence of slice leakage and noise amplification.

Bayesian image analysis in Fourier space for Medical Imaging
John Kornak, PhD (University of California, San Francisco)

Data from magnetic resonance imaging (MRI) modalities are acquired in terms of spatial frequencies, and are therefore representable in Fourier or k-space. The data are then typically inverse-Fourier-transformed, and the magnitude of the complex signal at each location is extracted to display an image. Bayesian image analysis can improve image quality, by balancing a priori expectations of image characteristics, with a model for the noise process via Bayes Theorem. In contrast to conventional Bayesian image analysis, we here preform Bayesian image analysis in Fourier space (BIFS), i.e., directly applied to the raw MRI data in k-space. By specifying the Bayesian model in Fourier space, spatially correlated priors, that are relatively difficult to model and compute in conventional image space, can be efficiently modeled as a set of independent processes across Fourier space; the priors in Fourier space are modeled as independent, but tied together by defining a function over Fourier space for the parameters. We will describe the BIFS modeling approach specifically for MRI, and demonstrate benefits to model specification, posterior image properties, and computational efficiency.

Bayesian Spatial Modeling via Kernel Convolutions on Complex-Valued fMRI Signals

Cheng-Han Yu, PhD (University of California, Santa Cruz)

We propose a novel complex-valued spatial model via kernel convolution (KC-S) and develop a MCMC algorithm to detect brain activation at the voxel level. The implementation of this model is done in a computationally efficient way by parcellating the voxel-based image into several spatial regions also by kernel-based dimension reduction. In general, spatial models encourage activation clustering and avoid falsely labeled isolated activated voxels. We show that the kernel-based model produces a more reasonable posterior probability map than the model proposed by Bezener et al. (2017) that use Gaussian processes for the region underlying spatial random effect. We then compare detection performance of the spatial models with the performance of the non-spatial model proposed in Yu et al. (2017). We find that when magnitude data are used, our kernel-based model much improves sensitivity especially when signal-to-noise ratio (SNR) is high and contrast-to-noise ratio (CNR) is low. When complex-valued data are used, the model without explicit spatial structure could perform as good as spatial models. This implies that complex-valued data may already include spatial information through phase angles that shows directions of complex-valued signals. As a result, without a sophisticated spatiotemporal model, using entire complex-valued data is recommended.

11:30 a.m. – 12:30 p.m.
Collaborative Case Study: Quantitative Immunohistochemistry Biomarkers based on Tissue Microarray Images
Organizer: Inna Chervoneva, PhD

Quantitative immunohistochemistry biomarkers for precision oncology
Hallgeir Rui, MD, PhD (Medical College of Wisconsin)

Protein markers in histological sections of malignant tumors hold great promise as predictors of response to therapy and prognosis. Immunohistochemistry (IHC) allows assessment of protein levels, subcellular localization, and degree of post-translational modifications such as phosphorylation. Historically, pathologists have visually estimated tumor marker levels in qualitative or at best semi-quantitative manner. Improvements in hardware, software and computational capacities are facilitating more objective digital marker quantification. Image analysis platforms now produce excellent marker data in tumor tissues. Tumor marker levels are typically computed as a single value, Mean Signal Intensity (MSI), within the tissue region of interest, typically the cancer cell region. Meanwhile, automated digital analyses allow capturing marker levels in each individual cell or subcellular compartment, but cell-level data remains largely underutilized. There is a need to develop statistical methods for discovery of new IHC biomarkers based on spatial localization of protein expression in cancer cells.

Spatial statistics approach to develop novel protein cancer biomarkers
Inna Chervoneva, PhD (Thomas Jefferson University)

We propose a new approach for developing IHC biomarkers using the information on spatial distribution of cellular signal intensity (CSI) of protein expression across the cancer cell population. We view protein expression levels as marks for the marked point process of cancer cells in the tumor tissue and develop spatial index predictors of clinical outcomes based on nonparametric spatial statistics describing the relationship between marks (protein expression) and points (cancer cell locations). The utility of the new spatial index IHC biomarkers is investigated and compared to the standard MSI predictors using the protein expressions in tissue microarrays (TMAs) incorporating tumor tissues from 2,000+ breast cancer patients. The new approach provides new insight into the standard breast cancer protein biomarkers and identifies novel protein biomarkers that do not have a prognostic value if only the mean signal intensity is considered.

1:45 p.m. – 3:35 p.m.

(Brief Break 2:45 p.m. – 2:55 p.m.)

Recent Advances in Modeling Large-Scale Imaging Data

Organizer: Zoe Zhang, PhD

Statistical modeling of brain networks using multimodal neuroimaging analysis
Ying Guo, PhD (Emory University)

In recent years, longitudinal neuroimaging study has become increasingly popular in neuroscience research to investigate disease-related changes in brain functions, to study neurodevelopment or to evaluate treatment effects on neural processing. One of the goals in longitudinal imaging analysis is to study temporal changes in brain functional networks and its association with subjects' clinical or demographic covariates. In this presentation, we propose a longitudinal independent component model (L-ICA) which provides a formal statistical modeling framework for extending ICA to longitudinal setting. By incorporating subject-specific random effects and visit-specific covariate effects, L-ICA is able to provide more accurate estimates of changes in brain functional networks on both the population- and individual-level, borrow information across repeated scans within the same subject to increase statistical power in detecting covariate effects on temporal changes in the networks, and allow for model-based prediction for changes in brain networks related to disease progression, treatment or
neurodevelopment. We develop EM algorithms to obtain maximum likelihood estimates of LICA and propose a statistical inference procedure for testing covariate effects in L-ICA. We demonstrate the performance of the proposed method via simulation studies and further illustrate its application to real-world longitudinal fMRI data.

NPBayes-fMRI: Nonparametric Bayesian General Linear Models for Single- and Multi-Subject fMRI Data
Jeong Hwan Kook (Rice University)

We present a Bayesian nonparametric regression model for the analysis of multiple-subject functional magnetic resonance imaging (fMRI) data. Our goal is to provide a joint analytical framework that allows the detection of regions of the brain that activate in response to a stimulus, while simultaneously taking into account the association, or clustering, of spatially remote voxels within and across subjects. The model incorporates information on both the spatial and temporal correlation structures of the data. It also allows for voxeldependent and subject-specific parameters. All methods are implemented in a Matlab userfriendly GUI.

Efficient semi-parametric regression for longitudinal data with regularized estimation of error covariance function
Chunming Zhang, PhD (University of Wisconsin-Madison)
Improving estimation efficiency for regression coefficients is an important issue in the analysis of longitudinal data, which involves estimating the covariance matrix of errors. But challenges arise in estimating the covariance matrix of longitudinal data collected at irregular or unbalanced time points. In this paper, we develop a regularization method for estimating the covariance function and a stepwise procedure for estimating the parametric components efficiently in the varying-coefficient partially linear model. This procedure is also applicable to the varying-coefficient temporal mixed effects model. Our method utilizes the structure of the covariance function and thus has faster rates of convergence in estimating the covariance functions and outperforms the existing approaches in simulation studies. This procedure is easy to implement and its numerical performance is investigated using both simulated and real data.

Sparse Multivariate Mediation and Moderated Mediation Analysis
Seonjoo Lee, PhD (Columbia University)

In investigating mechanisms underlying neurological or psychiatric disorders, we often have high-dimensional multivariate mediators. For example, in studying normal aging or pathology related cognitive decline, scientists want to evaluate whether the effect of aging or pathology on cognition is mediated by neural substrates measured by magnetic resonance imaging (MRI). To test the mediation hypothesis, mediators are typically summarized to couple of pre-identified regions of interests (ROIs) potentially resulting in information loss and bias in scientific findings. To fully utilize collected high- dimensional mediators, we propose a sparse multiple mediation method for high-dimensional brain image mediators. We form a sparse structural equation model and simultaneously select a parsimonious model by regularizing each mediation path instead of relying on a few pre-defined variables. For highly correlated mediators, principal component analysis is employed for dimension reduction and the proposed algorithm finds the relevant principal components that mediate the effect of the predictor on the outcome. Also we will discuss extension to moderated mediation analysis. Cortical thickness data from Reference Ability Neural Networks (RANN) and Alzheimer's Disease Neuroimaging Initiative (ADNI) were analyzed.

Posters

A Matlab Toolbox for Multivariate Analysis of Brain Networks

Mohsen Bahrami, Paul J. Laurienti, Sean L. Simpson*

Complex brain networks are believed to underlie behavioral and cognitive functions. Graph theoretical methods have extensively been used in studying such networks, and have provided promising results. However, development of multivariate frameworks that relate brain networks to phenotypic characteristics and draw statistical inference for such relationships, especially in evaluating whole-brain networks, has lagged behind. We have developed a freely-available Matlab toolbox, with a friendly graphical user interface (GUI), that bridges such important gaps between brain network analyses and statistical inference. While the toolbox is developed in Matlab, it uses SAS, R, or Python (depending on software availability) to perform the statistical modeling. The toolbox utilizes a mixed-effects regression framework that allows assessing brain network differences between study populations as well as assessing the effects of covariates of interest, such as disease phenotypes and risk factors, on brain connections in global (i.e., whole-brain) and local (i.e., subnetworks) brain networks. This toolbox can be used to evaluate various neuroimaging data types such as fMRI, EEG, MEG, PET, and DTI, which makes it useful for a wide range of studies evaluating the structure and function of brain networks. A k-means clustering-based data reduction method has also been implemented in this toolbox to help with model convergence and substantially reduce modeling time when large datasets are used.

Using the brain connectivity information under repeated measures design to find the association between cortical thickness and HIV disease

Damian Brzyski, Marta Karas, Beau Ances, Joaquin Goni, Timothy W. Randolph, Jaroslaw Harezlak*

One of the challenging problems in the brain imaging research is a principled incorporation of information from different imaging modalities in association studies. Frequently, data from each modality is analyzed separately using, for instance, dimensionality reduction tech-niques, which results in a loss of mutual information. Another important problem to address is incorporation of correlations among observations arising from repeated measurements on the same subject in a longitudinal study. We propose a novel regularization method, rePEER (repeated Partially Empirical Eigenvectors for Regression) to estimate the association be-tween the brain structure features and a scalar outcome. Our approach employs the external information about the brain connectivity (structural and functional) and takes into account the repeated measures designs. The method we propose is formulated as a penalized convex optimization problem. We address both theoretical and computational issues, such as the selection of tuning parameters which impacts the final estimate. We also show that utilizing the additional information significantly increases the estimation accuracy. We evaluated the performance of rePEER in simulation studies and applied it to analyze the association be-tween cortical thickness and HIV-related outcomes in the group of HIV-positive individuals.

An automated probabilistic algorithm for the detection of central vein sign in multiple sclerosis

Jordan D. Dworkin, Pascal Sati, Andrew Solomon, Dzung L. Pham, Richard Watts, Melissa L. Martin, Daniel Ontaneda, Matthew K. Schindler, Daniel S. Reich, Russell T. Shinohara*

Background: Central vein sign (CVS) is a promising diagnostic biomarker for multiple sclerosis (MS). Recent studies have demonstrated that patients with MS have higher proportions of CVS in white matter lesions compared to people with diseases that mimic MS on magnetic resonance imaging (MRI). However, the utility of the CVS biomarker is limited by inter-rater differences in the adjudication of CVS, as well as the time burden required for the determination of CVS for each lesion in a patient's scan.

Methods: The current study develops an automated technique for the detection of CVS in white matter lesions. The method is probabilistic in nature, allows for site-specific lesion segmentation methods, and has the potential to be

robust to inter-site variability. The proposed algorithm is tested on 40 participants at the University of Vermont; 20 patients with MS, 20 patients without.

Results: Using the automated technique, patients with MS were found to have significantly higher scores on the CVS biomarker than patients without MS (MMS = 0.55, Mnon-MS = 0.31, $p < 0.001$). The biomarker was also found to show strong discriminative ability between MS and non-MS patients, with an AUC value of 0.88.

Conclusion: The current study presents the first fully automated method for detecting central vein sign in white matter lesions, and demonstrates its promising performance in a sample of MS and non-MS patients.

Longitudinal Functional Principal Component Analysis in Structural MRI for Multiple Sclerosis

Menghan Hu, Ani Eloyan*

This project is motivated by a longitudinal MRI study of Multiple Sclerosis patients, where the objective was to study the progression of the MS disease by studying the intensity profiles of lesions. Understanding the dynamic behavior of the underlying trajectories is potentially useful for determining the stage of disease at the time of observation and predicting outcomes at future visits. We model the longitudinal MRI data as discrete observations from a functional process over time and apply functional data analytic methods in our study. The model also accounts for the hierarchical structure of the data where the variability of the functional data can be decomposed into three levels. MS patients develop lesions at seemingly random time points at various locations, hence the support of the data is unbalanced in both space and time. We group the data by their spatial characteristics and align the trajectories by time. To further reduce the computational complexity, we use principal component bases for the functional processes.

Predicting Natalizumab Effectiveness in Multiple Sclerosis Patients

Melissa Martin

Multiple sclerosis (MS) is a demyelinating central nervous system disease characterized by the incidence of lesions and the loss of brain volume (atrophy). In this study, we focused on a relatively new biological therapy, natalizumab, which reduces the migration of inflammatory immune cells into the brain, and has been shown to decrease the number of new lesions patients develop. Our goal was to predict how long a patient will continue natalizumab treatment based on lesion count, as well as changes in lesion volume and ventricular volume, using retrospective analyses of magnetic resonance imaging (MRI).

The dataset we worked with contained a sample of 35 MS patients who underwent MRI at three time points at the Hospital of the University of Pennsylvania. We performed a survival analysis to predict the number of days a patient stays on natalizumab treatment before stopping for ineffectiveness. The data indicate that patients who, before starting treatment, have more incident lesions or less atrophy tend to stay on natalizumab longer. Future investigation is needed to validate these exploratory findings.

Analytic white matter tractography and compositional distance based summarization of white matter brain structures

Wendy Meiring, Matthew Cieslak, Tegan Brennan, Subhash Suri, Scott T. Grafton*

We present an analytic (simulation free) method for calculating white matter transition probabilities between neighboring brain voxels based on DSI structural MRI, including phantom evaluation, human in-vivo studies, and Voxel Graph tractography illustrations. We also present two new voxel-wise univariate summary measures based on compositional data distances to describe features in white matter, including highlighting complex brain regions with splitting/fanning/kissing fibers. Our new summaries provide valuable additional insights beyond commonly used voxel-wise white matter descriptors such as fractional anisotropy (FA).

Weighted classification models with regularization to improve the reproducibility of neuroimaging signature selection

Xin Niu, Fengqing (Zoe) Zhang*

Multimodal neuroimaging has emerged as a promising approach to examine the imaging signatures of complex neurological disorders such as Alzheimer's disease and Parkinson's disease. Selection of the neuroimaging signatures is helpful for early disease diagnosis and assessment of disease progression. Combing features derived from different neuroimaging modalities (e.g., MRI, DTI, fMRI) is likely to improve our understanding of the structure and function of the brain by utilizing complementary physical and physiological sensitivities. However, the large number of derived features from different modalities poses great challenges for statistical variable selection. It is of critical importance that the selected neuroimaging features are robust and consistent across subjects. For example, cross-validation (CV) is commonly used to evaluate the performance of machine learning models. However, few studies examined the feature selection consistency across different folds of CV. In the current study, we propose a new way to quantify the reproducibility of neuroimaging signatures selected by classification models. Specifically, we incorporate the coefficient of variation of the mean difference between groups across 500 bootstrap samples as penalty weights in classification models with regularization (e.g., LASSO).

To evaluate our proposed models, we used multimodal brain imaging data including T1-weighted imaging, diffusion tenser imaging (DTI), and resting state functional brain imaging (rs-fMRI) from the PNC (Philadelphia Neurodevelopmental Cohort) dataset. We selected 91 subjects, among which 59 (43 females) were PTSD patients (PTSD group), 32 (8 females) were healthy controls who experienced traumatic events without long-lasting PTSD symptoms (trauma group). Gray matter volume (GMV), fractional anisotropy (FA) and mean diffusivity (MD) were extracted from the MRI and DTI data, respectively. Functional connectivity (FC) features were extracted from the rs-fMRI data. GMV features were averaged based on the brain regions in the Brainnetome atlases. DTI features were averaged based on the John Hopkins white matter atlas.

We compared the classification results of our proposed model with the LASSO model (Table 1). In addition, we quantified the consistency of selected features across cross-validation (CV) using R-index defined by the following formula:

$$R = \frac{\left|N - \frac{k}{2}\right| - \left(\frac{k}{2} - \left\lfloor\frac{k}{2}\right\rfloor\right)}{\left\lfloor\frac{k}{2}\right\rfloor}$$

Where N is the number of times a feature is selected in CV while K is the number of folds in CV. This index is only computed for features that are selected at least once in CV.

Our proposed model was effective in differentiating trauma and PTSD groups while obtaining more consistent selected features. Using our proposed model with FA features, a classification accuracy of 76.02% was achieved and the FA value in the left superior cerebellar peduncle was consistently selected across all folds. This finding is consistent with previous studies showing a close relation between left cerebellum and PTSD.

Table 1 Comparison of classification results between LASSO and our method for differentiating PTSD and trauma group.

	Standard LASSO			Our Method		
Features	Accuracy	SD	R-index	Accuracy	SD	R-index
VBM	0.6392	0.1318	0.4020	0.7158	0.1430	0.4643
FA	0.7158	0.0941	0.3043	0.7602	0.1256	0.5455
MD	0.6713	0.1658	0.3750	0.7047	0.1640	0.4444
FC	0.6275	0.1118	0.5000	0.6825	0.1385	0.4394
Multimodal	0.6058	0.2528	0.4118	0.6936	0.0723	0.4783

Intensity Normalization of T1-w MRI of Patients with Glioblastomas as a Preprocessing Tool to Improve Subsequent Radiomic Features

Abdhi Sarkar, Spyridon Bakas, Sung Min Ha, Christos Davatzikos, Russell T. Shinohara*

For patients with Glioblastomas multiple MR images are acquired of up to 6 different modalities. To improve automated segmentation methods or extract radiomic features to study the recurrence or survival of the patients, images across subjects need to be harmonized due to the well-known fact that MR images are acquired in arbitrary units. We propose a technique that uses tissue-based histogram matching by masking out tumor related Edema seen as hyper intense on a FLAIR image. Using FSL FAST, a tissue segmentation technique, we match each patient image to a single healthy control or template image based on only a small sample of voxels that have the highest probability of belonging to a certain tissue class. After this process, we observe a significant improvement in the alignment of histograms across the entire brain image for all subjects and improved contrast in the whole brain images as well.

Structural and functional asymmetry of medial temporal subregions in unilateral temporal lobe epilepsy: a 7T MRI study

Preya Shah, Danielle S. Bassett, Laura E.M. Wisse, John A. Detre, Joel M. Stein, Paul A. Yushkevich, Russell T. Shinohara, Mark A. Elliott, Sandhitsu R. Das, Kathryn A. Davis*

Mesial temporal lobe epilepsy (TLE) is a common neurological disorder affecting the hippocampus and surrounding medial temporal lobe (MTL). While prior studies have analyzed whole-brain network distortions in TLE patients, the functional network architecture of the MTL at the subregion level has not been examined. In this study, we utilized high-resolution 7T T2-MRI and resting-state BOLD-fMRI to characterize volumetric asymmetry and functional network asymmetry of MTL subregions in unilateral medically-refractory TLE patients and healthy controls. We subdivided the TLE group into mesial temporal sclerosis patients (TLE-MTS) and MRI-negative nonlesional patients (TLE-NL). Using an automated multi-atlas segmentation pipeline, ten MTL subregions were delineated per hemisphere for each subject. TLE-MTS demonstrated volumetric asymmetry corresponding to decreased volumes ipsilaterally in the whole hippocampus and multiple individual MTL subregions ($p < 0.05$, two-tailed two-sample permutation test). In contrast, TLE-NL demonstrated functional network asymmetry corresponding to increased connectivity ipsilaterally in the whole hippocampus and CA1 subfield ($p < 0.05$, two-tailed two-sample permutation test). Findings did not change after regressing out effects of age and sex. A logistic regression model of TLE groups vs. controls revealed that volumetric asymmetry could be used to effectively classify TLE-MTS (AUC = 0.94), but not TLE-NL (AUC = 0.42). In contrast, functional network asymmetry could be used to classify TLE-NL (AUC = 0.84) more effectively than TLE-MTS (AUC = 0.63). Our findings provide initial evidence that functional neuroimaging-based network properties within the MTL can serve as useful biomarkers for TLE cases in which conventional imaging and volumetric analysis is insufficient. High resolution MRI has potential to improve localization of underlying brain network disruptions in TLE patients who are candidates for surgical resection.

A Dual Modeling Approach to Automatic Segmentation of Cerebral T2 Hyperintensities and T1 Black Holes in Multiple Sclerosis

Alessandra M. Valcarcel, Kristin A. Linn, Fariha Khalid, Simon N. Vandekar, Shahamat Tauhid, Theodore D. Satterthwaite, John Muschelli, Rohit Bakshi, Russell T. Shinohara*

Background and Purpose: Magnetic resonance imaging (MRI) is crucial for in vivo detection and characterization of white matter lesions (WML) in multiple sclerosis (MS). The most widely established MRI outcome measure is the volume of hyperintense lesions on T2-weighted images (T2L). Unfortunately, T2L volume is non-specific for the level of tissue destruction and shows a weak relationship to clinical status. Consequently, some researchers have focused on quantification of the volume of hypointense lesions on T1-weighted images (T1L) ("black holes") to provide more specificity for axonal loss and a closer link to neurologic disability. The technical difficulty of T1L segmentation has led investigators to rely primarily on time-consuming manual methods which are prone to inter- and intra-rater variability. This study aimed to develop an automatic T1L segmentation algorithm, adapted from an automatic T2L method. First, to ensure the method performed well on the images of interest, the automated T2L segmentation was compared with gold standard manual segmentations. The pipeline was then adapted to automated T1L segmentation, with validation versus gold standard manual segmentations and clinical-MRI correlations.

Materials and Methods: High-resolution 3D T1-w, T2-w, and FLAIR sequences were acquired from 40 MS subjects on a Siemens 3T Skyra unit at the Brigham and Women's Hospital (Boston, MA). Trained observers under the supervision of an experienced observer manually segmented T2L and T1L. MIMoSA, an automated segmentation algorithm, was then employed which has previously shown competitive performance for segmenting T2L. MIMoSA utilizes complementary imaging modalities to emphasize different tissue properties, which can help identify and characterize interrelated features of lesions, in a local logistic regression to model the probability that any voxel is part of a lesion.

Results: Using bootstrap cross-validation, MIMoSA proved to be a robust method that can be applied to segment both T2L and T1L. For T2L, results yield a Sørensen-Dice coefficient (DSC) of 0.6, partial AUC (pAUC) up to 1% false positive rate of 0.69. For T1L, performance results in 0.48 DSC and 0.63 pAUC. Correlation between manual and automatic volumes with EDSS were recorded. Manual volumes resulted in an EDSS correlation of 0.32 with T1L, 0.34 with T2L, 0.33 with T1L/T2L. MIMoSA volumes showed similar results, with correlations of 0.34 for T1L, 0.34 for T2L, and 0.28.

Conclusions: MIMoSA is a fully automated segmentation algorithm. Though originally designed to segment T2L, MIMoSA is able to segment T1 black holes with both sensitivity and specificity in patients with MS.

Powerful Permutation Tests for Neuroimaging using Voxel-wise Transformations

Simon N. Vandekar, Theodore D. Satterthwaite, Adon Rosen, Rastko Ciric, David R. Roalf, Kosha Ruparel, Ruben C. Gur, Raquel E. Gur, Russell T. Shinohara*

Typical statistical methods in neuroimaging result in hundreds of thousands of tests performed in an image, followed by the use of a multiple testing procedure (MTP) to control the family-wise error rate (FWER). Recent studies have demonstrated that widely used MTP procedures yield anticonservative FWERs. The permutation MTP is among few procedures shown to reliably control the number of false positives at the specified probability. Voxel-wise permutation tests work by randomly permuting the imaging data and using the distribution of the maximum value of the test statistic across all voxels in the image to compute adjusted p-values. While this procedure has intuitive appeal, anecdotally many investigators have noted it lacks power. We demonstrate that the procedure lacks power because neuroimaging data have voxels with heavy skew near the edge of the brain. These voxels cause the distribution of the maximum across the image to be heavily inflated. As a solution we apply the Yeo-Johnson transformation prior to permutation testing. The transformation yields a statistical image where all the voxels have approximately the same distribution and improves the power of the test.

Generative discriminative regression for neuroimaging analysis

Erdem Varol

We propose a general framework for obtaining interpretable multivariate discriminative models that allow efficient statistical inference for neuroimaging analysis. The framework, termed generative discriminative regression (GDR), augments discriminative models with a generative regularization term. We demonstrate that the proposed formulation can be optimized in closed form and in dual space, allowing efficient computation for high dimensional neuroimaging datasets. Furthermore, we provide an analytic estimation of the null distribution of the model parameters, which enables efficient statistical inference and p-value computation without the need for permutation testing. We compared the proposed method with both purely generative and discriminative learning methods in two large structural magnetic resonance imaging (sMRI) datasets of Alzheimer's disease (AD) (n=415) and Schizophrenia (n=853). Using the AD dataset, we demonstrated the ability of GDR to robustly handle confounding variations. Using Schizophrenia dataset, we demonstrated the ability of GDR to handle multi-site studies. Taken together, the results underline the potential of the proposed approach for neuroimaging analyses.

Quantifying Site and Scanner variability in structural MRIs

Julia Wrobel

Magnetic resonance imaging is used to detect structural changes in the brain. Quantitative analysis in MRI studies can be a challenge due to arbitrary intensity units that differ across scans, and intensity normalization addresses this problem. We present an intensity normalization technique that uses non-linear alignment of CDFs of the voxel intensifies. Our goal is to quantify site and scanner variability in biologically stable images and determine how well our method removes these effects. We find that unnormalized images are highly variable across site and scanner type, and our method effectively removing this variability by warping the intensities. Ability to predict intensity normalization results for a scan taken on an existing subject at a new site is evaluated using cross-validation.

Poster Numbers & Titles

1

A Matlab Toolbox for Multivariate Analysis of Brain Networks

Mohsen Bahrami, Paul J. Laurienti, Sean L. Simpson*

2

Using the brain connectivity information under repeated measures design to find the association between cortical thickness and HIV disease

Damian Brzyski, Marta Karas, Beau Ances, Joaquin Goni, Timothy W. Randolph, Jaroslaw Harezlak*

3

An automated probabilistic algorithm for the detection of central vein sign in multiple sclerosis

Jordan D. Dworkin, Pascal Sati, Andrew Solomon, Dzung L. Pham, Richard Watts, Melissa L. Martin, Daniel Ontaneda, Matthew K. Schindler, Daniel S. Reich, Russell T. Shinohara*

4

Longitudinal Functional Principal Component Analysis in Structural MRI for Multiple Sclerosis

Menghan Hu, Ani Eloyan*

5

Predicting Natalizumab Effectiveness in Multiple Sclerosis Patients

Melissa Martin

6

Analytic white matter tractography and compositional distance based summarization of white matter brain structures

Wendy Meiring, Matthew Cieslak, Tegan Brennan, Subhash Suri, Scott T. Grafton*

7

Weighted classification models with regularization to improve the reproducibility of neuroimaging signature selection

Xin Niu, Fengqing (Zoe) Zhang*

8

Intensity Normalization of T1-w MRI of Patients with Glioblastomas as a Preprocessing Tool to Improve Subsequent Radiomic Features

Abdhi Sarkar, Spyridon Bakas, Sung Min Ha, Christos Davatzikos, Russell T. Shinohara*

9

Structural and functional asymmetry of medial temporal subregions in unilateral temporal lobe epilepsy: a 7T MRI study

Preya Shah, Danielle S. Bassett, Laura E.M. Wisse, John A. Detre, Joel M. Stein, Paul A. Yushkevich, Russell T. Shinohara, Mark A. Elliott, Sandhitsu R. Das, Kathryn A. Davis*

10

A Dual Modeling Approach to Automatic Segmentation of Cerebral T2 Hyperintensities and T1 Black Holes in Multiple Sclerosis

Alessandra M. Valcarcel, Kristin A. Linn, Fariha Khalid, Simon N. Vandekar, Shahamat Tauhid, Theodore D. Satterthwaite, John Muschelli, Rohit Bakshi, Russell T. Shinohara*

11

Powerful Permutation Tests for Neuroimaging using Voxel-wise Transformations

Simon N. Vandekar, Theodore D. Satterthwaite, Adon Rosen, Rastko Ciric, David R. Roalf, Kosha Ruparel, Ruben C. Gur, Raquel E. Gur, Russell T. Shinohara*

12

Generative discriminative regression for neuroimaging analysis

Erdem Varol

13

Quantifying Site and Scanner variability in structural MRIs

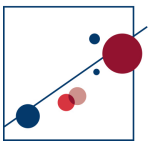
Julia Wrobel

Organizing Committee

Dan Kessler, *University of Michigan*
Kristin Linn, *University of Pennsylvania*
Amanda Mejia, *Indiana University*
John Muschelli, *Johns Hopkins University*
Inna Chervoneva, *Thomas Jefferson University*
Zoe Zhang, *Drexel University*
Lucy Robinson, *Drexel University*
Taki Shinohara, *University of Pennsylvania*
Haochang Shou, *University of Pennsylvania*

The organizing committee thanks **Marissa Fox** and **Meghan LeCates,** whose efforts made this conference possible.

This conference was co-sponsored by the Center for Clinical Epidemiology and Biostatistics at the University of Pennsylvania and the American Statistical Association.



CENTER FOR
CLINICAL EPIDEMIOLOGY
AND BIOSTATISTICS

ASA
AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics®