Emerging Issues in the Analysis of Longitudinal Data

August 16 – August 21, 2009

Banff, Alberta, Canada

MEALS

*Breakfast (Buffet): 7:00–9:00 am, Sally Borden Building, Monday–Friday
*Lunch (Buffet): 11:30 am–1:30 pm, Sally Borden Building, Monday–Friday
*Dinner (Buffet): 5:30–7:30 pm, Sally Borden Building, Sunday–Thursday

*Please remember to scan your meal card at the host/hostess station in the dining room for each meal.

MEETING ROOMS

All lectures will be held in Max Bell 159 (Max Bell Building accessible by walkway on 2nd floor of Corbett Hall) at the Banff Centre. LCD projector, overhead projectors and blackboards are available for presentations.

Each talk is roughly 30-minute. The talks are arranged according to the themes for each day, which are based on the proposal/objectives of the workshop.

Please note that the meeting space designated for BIRS is the lower level of Max Bell, Rooms 155–159. Please respect that all other space has been contracted to other Banff Centre guests, including any Food and Beverage in those areas.

SCHEDULE

Sunday, August 16. Check-in

16:00 –	Check-in begins (Front Desk - Professional Development Centre - open 24 hours) Lecture rooms available after 16:00 (if desired)
17:30 – 19:30	Buffet Dinner, Sally Borden Building
20:00 –	Informal gathering in 2nd floor lounge, Corbett Hall (if desired) Beverages and small assortment of snacks available on a cash honour-system.

Monday, August 17, Missing Data and Measurement Error

7:00 − 8:45	Breakfast
8:45 − 9:00	Introduction and Welcome, Max Bell 159
	– BIRS Station Manager
	– Charmaine Dean, Former President of SSC & Co-organizer of the workshop
9:00 − 10:30	**Missing Data (I)**
	Chair: Charmaine Dean, Simon Fraser University
9:00	Roderick Little, University of Michigan
	Missing values in longitudinal data analysis: some considerations
9:30	Joan Hu, Simon Fraser University
	Estimation under Cox proportional hazards model with nonrandom covariate missing
10:00	Annie Qu, University of Illinois at Urbana-Champaign
	Highly efficient aggregate unbiased estimating functions approach for correlated data with missing at random
10:30 − 11:00	**Coffee Break, 2nd floor lounge, Corbett Hall**
11:00 − 12:00	**Missing Data (II) − Multiple Imputation**
	Chair: Charmaine Dean, Simon Fraser University
11:00	Mike Kenward, University of London
	Using multiple imputation to obtain doubly robust estimators
11:30	James Carpenter, University of London
	Multilevel multiple imputation with mixed response types
12:00 − 13:30	**Lunch**
13:30 − 14:30	**Marginal Models for Incomplete Data & Misspecification**
	Chair: Grace Yi, University of Waterloo
13:30	Richard Cook, University of Waterloo
	Marginal methods for estimating treatment effects in cluster randomized longitudinal studies with incomplete responses and non-compliance
14:00	Baojiang Chen, University of Washington
	Weighted generalized estimating functions for incomplete longitudinal response and covariate data
14:30 − 15:00	**Coffee Break, 2nd floor lounge, Corbett Hall**
15:00 − 16:00	**Discussion Session**: *Sensitivity analysis for longitudinal data with dropout, what should we be doing, and how?*
	Organizer & Chairs: James Carpenter & Mike Kenward, University of London
	Recorder: Baojiang Chen, University of Washington
16:00 −	**Guided Tour of The Banff Centre; meet in the 2nd floor lounge, Corbett Hall**
17:30 − 19:30	**Dinner**

Tuesday, August 18. Functional Data, Mixed Models, Estimating Equations

7:00–9:00	Breakfast
9:00 − 10:30	**Semiparametric (Functional) Methods (I)**
	Chair: Xihong Lin, Harvard University

9:00   Raymond Carroll, Texas A&M University
         Efficient inference in additive models with repeated measures

9:30   Liang Hua, University of Rochester
         Variable selection for semiparametric models with measurement errors

10:00   Lu Wang, University of Michigan
         Nonparametric regression in longitudinal studies with dropout at random

10:30 − 11:00   **Coffee Break, 2nd floor lounge, Corbett Hall**

11:00 − 12:00   **Semiparametric/Nonparametric Methods**
         Chair: Xihong Lin, Harvard University

11:00   Naisyin Wang, Texas A&M University and University of Michigan
         Functional latent feature regression models for data with longitudinal covariate process

11:30   Wenqing He, University of Western Ontario
         Local linear regression on clustered censored data

12:00 − 13:30   **Lunch**

13:30 − 14:30   **Mixed Effects Models**
         Chair: John Neuhaus, University of California at San Francisco

13:30   Charles McCulloch, University of California at San Francisco
         Estimation efficiency in generalized linear mixed models under
           misspecified random effects distributions

14:00   Daniel Li, University of Manitoba
         Second-order least squares estimation of mixed effects models

14:30 − 15:00   **Coffee Break, 2nd floor lounge, Corbett Hall**

15:00 − 16:00   **Estimating Equations**
         Chair: John Neuhaus, University of California at San Francisco

15:00   Peter Song, University of Michigan
         Analyzing unequally spaced longitudinal data with quadratic inference functions

15:30   Youngjo Lee, Seoul National University
         Hierarchical generalized linear models (HGLMs) and variable selection

16:00 −   **Group Photo; meet on the front steps of Corbett Hall**

17:30 − 19:30   **Dinner**

Wednesday, August 19, Joint Models

7:00–9:00   Breakfast
9:00 − 10:30   **Joint Models**
                   Chair: Wei Liu, York University

9:00   Joseph Ibrahim, University of North Carolina at Chapel Hill
                      Local influence for joint models for longitudinal and survival data
9:30   Jeremy Taylor, University of Michigan
                      Using joint models for longitudinal and survival data to give individual predictions
10:00   Bin Nan, University of Michigan
                      Joint modeling of longitudinal and survival data when the event time is a covariate

10:30 − 11:00   **Coffee Break, 2nd floor lounge, Corbett Hall**

11:00 − 12:00   **Discussion Session**: *Longitudinal analysis using joint models*
                   Organizer & Chairs: Charmaine Dean, Simon Fraser University
                   Recorder: Wei Liu, York University

12:00 − 13:30   **Lunch**

13:30 −   **Free afternoon: sight-seeing**

17:30 − 19:30   **Dinner**

Thursday, August 20, Important Applications; Binary and Count Data

7:00 − 9:00	Breakfast
9:00 − 10:30	**New Methods for Important Applications**
	Chair: Liqun Wang, University of Manitoba
9:00	John Petkau, University of British Columbia
	Neutralizing antibodies and the efficacy of interferon Beta-1b in multiple sclerosis clinical trials
9:30	Andrea Rotnitzky, Harvard University
	Estimation and extrapolation of optimal dynamic treatment and testing strategies from observational longitudinal data
10:00	Tze-Leung Lai, Stanford University
	A dynamic empirical Bayes approach to econometric panel data via generalized linear mixed models
10:30 − 11:00	**Coffee Break, 2nd floor lounge, Corbett Hall**
11:00 − 12:00	**Binary and Count Data with Measurement Error**
	Chair: Liqun Wang, University of Manitoba
11:00	Taraneh Abarin, Samuel Lunenfeld Research Institute
	Instrumental variable approach to covariate measurement error in generalized linear models
11:30	Zhijian Chen, University of Waterloo
	A marginal method for correlated binary data with misclassified responses
12:00 − 13:30	**Lunch**
13:30 − 14:30	**Discussion Session**: *Random effects modeling of longitudinal data with excessive zeros*
	Organizer & Chairs: Renjun Ma, University of New Brunswick
	Recorder: Lang Wu, University of British Columbia
14:30 − 15:00	**Coffee Break, 2nd floor lounge, Corbett Hall**
15:00 − 16:00	**Modeling Complex Data**
	Chair: Jiayang Sun, Case Western Reserve University
15:00	Li Qin, Fred Hutchinson Cancer Center & University of Washington
	A registration-based functional linear model for post-ART viral loads in patients with primary HIV infection
15:30	Yang Zhao, University of Regina
	Likelihood method for regression models with data missing at random
16:00 −	**Informal gathering in groups to finalize summaries of discussions**
17:30 − 19:30	**Dinner**

Friday, August 21, Summaries of Discussions and Future Directions

7:00 − 8:30   Breakfast

8:30 − 9:30   **Discussion Session**: *Emerging issues of longitudinal data analysis*
Organizer & Chairs: Xihong Lin and Grace Yi

9:30 − 10:30   **Summaries From Discussion Sessions & Future Research Directions**
Organizers of the discussion sessions present summaries of the discussions
Chair: Lang Wu, University of British Columbia

10:30 − 11:00   **Coffee Break, 2nd floor lounge, Corbett Hall**

11:00 −   **Informal discussion. Preparation for departure. Checkout by 12 noon.**

11:30–13:30   Lunch

You are welcome to use the BIRS facilities (2nd Floor Lounge, Max Bell Meeting Rooms, Reading Room) until 3 pm on Friday, although participants are still required to checkout of the guest rooms by 12 noon.

DISCUSSIONS

Session I: **Monday, August 17, 3:00pm – 4:00pm**
Organizers: **James Carpenter** and **Mike Kenward**
Title: *Multilevel multiple imputation with mixed response types*

Outline of the discussion session:
1. 3:00pm – 3:05pm: Mike Kenward - Introduction
2. 3:05pm – 3:15pm: James Carpenter - Sensitivity analysis via multiple imputation: ideas and an example
3. 3:15pm – 3:30pm: Andrea Rotnitsky - Sensitivity analysis for the estimation of the effect of dynamic treatment regimes under flexible dynamic visit regimes
4. 3:30pm – 3:45pm: Joe Ibrahim - Local influence approaches for sensitivity analysis
5. 3:45pm – 4:00pm: Discussion, facilitated by Ray Carroll.

Session II: **Wednesday, August 19, 11:00am – 12:00pm**
Organizer: **Charmaine Dean**
Title: *Longitudinal analysis using joint models*

Outline of the discussion session:
1. Brief Overview
2. Illustrations of Applications
3. Challenges in Joint Modeling
4. Benefits and Drawbacks.
The discussion leader will provide introductory remarks to each of these sub-topics and audience participation will occur through the discussion of each of these four these areas. The discussion will also evolve depending on directions set by participants. We will close with a summary of major points.

Session III: **Thursday, August 20, 1:30pm - 2:30pm**
Organizer: **Renjun Ma**
Title: *Random effects modeling of longitudinal data with excessive zeros*

Outline of the discussion session:
1. various applications in different subject areas (data sets)
2. different approaches to random effects modeling of longitudinal data with excessive zeros in the literature (models, estimation methods and etc.). The relative advantages and limitations of these approaches
3. propose new approaches to random effects modeling of data with excessive zeros.
4. open the floor for discussions on further studies/collaborations.

Session IV: **Friday, August 21, 8:30am – 9:30am**
Organizer: **Xihong Lin** and **Grace Yi**
Title: *Emerging issues of longitudinal data analysis*

Outline of the discussion session:
1. analysis of high-dimensional data in longitudinal studies
2. measurement error problems
3. joint inference
4. semiparametric and nonparametric regression methods.

Emerging Issues in the Analysis of Longitudinal Data
August 16 – August 21, 2009

ABSTRACTS
(in alphabetic order by speaker surname)

Abarin, Taraneh (Samuel Lunenfeld Research Institute - Mount Sinai Hospital)

Title: *Instrumental Variable Approach to Covariate Measurement Error in Generalized Linear Models*

Abstract: We present the method of moments estimation for generalized linear measurement error models using the instrumental variable approach. The measurement error has a parametric distribution that is not necessarily normal, while the distributions of the unobserved covariates are nonparametric We also propose simulation-based estimators for the situation where the closed forms of the moments are not available. The proposed estimators are strongly consistent and asymptotically normally distributed under some regularity conditions. Finite sample performances of the estimators are investigated through simulation studies. Joint work with Liqun Wang

Carpenter, James (University of London)

Title: *Multilevel multiple imputation with mixed response types*

Abstract: We build upon the existing literature to formulate a class of models for multivariate mixtures of Gaussian, ordered or unordered categorical responses and continuous distributions that are not Gaussian, each of which can be defined at any level of a multilevel data hierarchy. We describe a MCMC algorithm for fitting such models. We show how this approach can be used to implement multilevel multiple imputation (assuming data are missing at random), and extended to allow imputation of missing data that is congenial/consistent with a complex multilevel model. We apply our approach to the analysis of the Health Behavior in School-aged Children (HBSC) study, an international survey conducted in collaboration with the WHO Regional Office for Europe.

Carroll, Raymond (Texas A&M University)

Title: *Efficient Inference in Additive Models With Repeated Measures*

Abstract: This talk considers the general problem where the data for an individual are repeated measures in the most general sense, with a parametric component and an additive nonparametric component. Using smooth backfitting, we derive efficient estimates of the parametric and nonparametric components, estimates that take the repeated measures design into account. In gene-environment interaction studies, it is often of interest to test for the main effects of genes (the parametric components) when there might be interactions with the environment (the nonparametric components). Rather than build complex models for the interactions, we use a Tukey-type 1-degree of freedom formulation that has the promise to improve power for testing whether there are any genetic effects. We derive a general profile-type score statistic and show how to implement it, which involves circumventing the need to solve an integral equation. One of the key features of this problem is that algorithms involve fitting misspecified models: we show that our methods, which involve a project, neatly overcome this issue.

Chen, Baojiang (University of Washington)

Title: *Weighted Generalized Estimating Functions for Incomplete Longitudinal Response and Covariate Data*

Abstract: Data from longitudinal studies often feature both missing responses and missing covariates. When the response is missing, the probability a particular covariate is missing is often higher, as a result of a positive association between the missingness for the response and covariates at each follow-up assessment. The impact of missing data in these settings depends on the frequency data are missing and the strength of the association between the missing data process and response process. In the setting of incomplete response and covariate data it is important to take the association between these processes into account when analyzing data. Inverse probability weighted generalized estimating equations offer one method for doing this and we develop these here. Empirical studies demonstrate that the consistent estimators arising from the proposed methods have very small empirical biases in moderate samples, and are more efficient than alternative methods which ignore the association between the missing data processes. Joint work with Grace Yi and Richard Cook.

Chen, Zhijian (Charlie) (University of Waterloo)

Title: A marginal method for correlated binary data with misclassified responses.

Abstract: Measurement error or misclassification has been a long standing concern in medical research and survey sampling. Much research has been directed to problems concerning error-prone covariates. There is relatively little work on measurement error or misclassification in the response variable (e.g., Carroll et al. 2006, Neuhaus 2002). In this talk we will focus the discussion on misclassification in clustered or longitudinal binary data. We propose a marginal analysis method to handle binary response which is subject to misclassification. Numerical studies will be presented to assess the performance of the proposed methods. Joint work with Grace Yi and Changbao Wu.

Cook, Richard (University of Waterloo)

Title: Marginal methods for estimating treatment effects in cluster randomized longitudinal studies with incomplete responses and non-compliance

Abstract: Cluster-randomized longitudinal studies are often designed to assess the effects of experimental behaviourial interventions. Examples include smoking cessation studies, family studies of dietary interventions, and studies of the effect of different programs for physical activity. Behavioural intervention studies frequently feature missing responses and individuals are often non-compliant. We describe methods for estimating the effect of interventions which account for these complications. Inverse weighted generalized estimating equations are used to address incomplete compliance data in a model for the compliance process. A "mean-score" approach is used to deal with the missing compliance data in the response model, since compliance status is incompletely observed. Simulation studies demonstrate that this method yields estimators with good frequency properties which are superior to those of standard alternative approaches. This is joint work with Baojiang Chen (Waterloo) and Grace Yi (Waterloo)

Dean, Charmaine/Zeng, Leilei (Simon Fraser University)

Title: Joint Analysis of Longitudinal Growth and Interval Censored Mortality Data

Abstract: Joint analysis of longitudinal and survival data has received considerable attention in the recent literature. We review methods developed for such joint analysis and develop a joint model for the analysis of longitudinal data monitoring the growth and survival of trees subject to various interventions in a designed experiment. Of interest is the development of methods for handling features of the data which are not common in considerations of joint analyses. Two main features are interval censoring of the survival response and the need to account for spatial effects jointly in the survival and longitudinal outcomes. We adopt linkages in random effects over multiple outcomes such as considered in Chi and Ibrahim (2006). To handle interval censoring in the survival context we propose to use functional approaches to smooth

the data. To account for covariate effects, we propose an iterative approach in a framework similar to an accelerated failure time formulation: covariate effects are used to back-transform the timescale for individual level data and then smoothing methods imposed on the back-transformed time-scales. Properties of the algorithmic scheme of iteration between estimation of covariate effects and imputation of the survival data using smoothing on back-transformed data will be developed in the context of the spatial joint analysis. We also discuss the conditions under which there are efficiency gains in joint analyses with regard determination of treatment effects. Joint work with Leilei Zeng and Terry Lee.

He, Wenqing (University of Western Ontario)

Title: Local Linear Regression on Clustered Censored Data

Abstract: Regression on clustered data is of considerable interest in the last decades. The response considered in those clustered regression models may be censored, and thus traditional methods have to be modified to incorporate censoring. In this talk, we present a local linear regression method for the estimation of the relationship between censored response and covariates by considering a transformation of the censored response. Simulation study is conducted to assess the performance of the proposed method, and a real example is employed as an illustration. This is joint work with Zhezhen Jin (Columbia University).

Hu, Joan (Simon Fraser University)

Title: Estimation under Cox Proportional Hazards Model with Nonrandom Covariate Missing

Abstract: Motivated by a study for disease control, we consider estimation under the Cox proportional hazards model using a set of right-censored survival times with missing covariates, where the covariate missing is not at random conditional on the observed data. We present a likelihood based estimation procedure with the available data supplemented with some readily available information. The medical study that motivated this research is used throughout the talk for illustration.

Ibrahim, Joe (University of North Carolina at Chapel Hill)

Title: Local Influence for Joint Models for Longitudinal and Survival Data

Abstract: We carry out an in-depth investigation of diagnostic measures for assessing the influence of observations and model misspecification for longitudinal models and joint models for longitudinal and survival data, allowing for the presence of missing data. We consider a local influence approach and examine various perturbation schemes for perturbing the models in this setting. We develop a perturbation manifold, and develop various local influence measures to identify influential points and test model misspecification. Simulation studies are conducted to evaluate our methods, and a real dataset is analyzed to illustrate the use of our methodology.

Kenward, Mike (University of London)

Title: Using Multiple Imputation to Obtain Doubly Robust Estimators

Abstract: In a recent contribution by Bang and Robins (Biometrics, 2005, 61, 962-973) it was shown how doubly robust estimators for monotone incomplete data problems can be obtained through a sequence of regressions, making such methods more easily accessible, although no explicit variance estimator was provided. It is shown in this talk how reformulation of this regression approach in a multiple imputation framework leads to very convenient route for calculating such estimators, while providing, at the same time, an explicit and easily calculable variance estimator. Further, it is shown how, by using the fully conditional specification for the imputation procedure, this approach to doubly robust estimation can extended to non-monotone missing value settings. This is a joint work with Rhian Daniel.

Lai, Tze Leung (Stanford University)

Title: A Dynamic Empirical Bayes Approach to Econometric Panel Data via Generalized Linear Mixed Models

Abstract: We first give a brief review of the literature on credibility rate-making in insurance and default modeling of corporate loans in finance, particularly on the econometric models used to analyze the associated panel data. We then propose a new, unified class of dynamic empirical Bayes models for these longitudinal data and their subject-matter applications. The advantages of these models and their connections to generalized linear mixed models are also discussed and illustrated.

Lee, Youngjo (Seoul National University)

Title: HGLM and variable selection

Abstract: Hierarchical Generalized Linear Models (HGLMs) provide a flexible and efficient framework for modeling non-Normal data when there are several sources of error variation. They extend the familiar generalized linear models (GLMs) to include additional random terms in the linear predictor. They include generalized linear mixed models (GLMMs) as a special case, but do not constrain the random terms to have a Normal distribution nor an identity link: e.g. if the basic GLM is a log-linear model (Poisson distribution and log link), an alternative assumption for the random terms might be a gamma distribution and log link. HGLMs thus bring a wide range of models together within a single framework. They also facilitate the joint modeling of mean and dispersion (of any random term). Main focus of this talk is to show how the HGLM can be used for variable selection. We show how LASSO and its extension can be obtained via random-effect models.

Li, Daniel (University of Manitoba)

Title: Second-order least squares estimation of mixed effects models

Abstract: The main approach to estimation of mixed effects models focuses on the likelihood methods which rely heavily on the normality assumption for random effects and regression random errors. In addition, they often have computational difficulties. We apply the second-order least squares method to estimate generalized linear mixed effects models where the distributions of the regression errors are nonparametric while those of random effects are parametric but not necessarily normal. These estimators are based on the first two marginal moments of the response variable. We present simulation studies of finite sample properties of the second-order least squares estimators and compare them with the maximum likelihood estimators. We also investigate the robustness of the estimators when the regression error and/or random effects distributions are misspecified. Joint work with Liqun Wang

Liang, Hua (University of Rochester)

Title: Variable selection for semiparametric models with measurement errors

Abstract: We explore variable selection for partially linear models when the covariates are measured with additive errors. We propose two classes of variable selection procedures, penalized least squares and penalized quantile regression, using the nonconvex penalized principle. The first procedure corrects the bias in the loss function caused by the measurement error by applying the so-called correction-for-attenuation approach, whereas the second procedure corrects the bias by using orthogonal residuals. The sampling properties for the two procedures are investigated. The rate of convergence and the asymptotic normality of the resulting estimates are established. We further demonstrate that, with proper choices of the penalty functions and the regularization parameter, the resulting estimates perform asymptotically as well as an oracle procedure. Choice of smoothing parameters is also discussed. Finite sample performance of the

proposed variable selection procedures is assessed by Monte Carlo simulation studies. We further illustrate the proposed procedures by an application.

Little, Roderick (University of Michigan)

Title: Missing values in longitudinal data analysis: some considerations

Abstract: Missing values are a common problem in longitudinal data sets, because of attrition, missed visits and other causes. Likelihood-based methods for handling this problem, based on parametric models for the data and missing-data mechanism, are discussed. Topics reviewed include (a) pros and cons of different forms of likelihood inference, specifically maximum likelihood, Bayes and multiple imputation; (b) robust estimation under the missing at random assumption based on penalized spline of propensity models; and (c) non-MAR models for the joint distribution of the repeated measures Y and the missing data indicators M. Concerning (c), I'll discuss the role of sensitivity analysis, and models formed by different factorizations of Y and M, including selection and pattern-mixture models, shared parameter models, and others.

McCulloch, Charles (University of California at San Francisco)

Title: Estimation efficiency in generalized linear mixed models under misspecified random effects distributions.

Abstract: Statistical models that include random effects are commonly used to analyze longitudinal and correlated data, often with the assumption that random effects follow a Gaussian distribution. Previous work has shown that incorrect specification of this distribution typically produces little bias in estimates of covariate effects and very modest inaccuracy in predicted random effects but few studies have assessed the effect of misspecification on standard errors and statistical tests. In this talk we examine the impact of a misspecified random effects distribution on estimation efficiency. We motivate this work using a variety of examples and then address the question using theoretical calculations and simulation. Joint work with John Neuhaus.

Nan, Bin (University of Michigan)

Title: Joint modeling of longitudinal and survival data when the event time is a covariate

Abstract: It is of great scientific interest to estimate the hormone profile, such as serum estradiol or follicle stimulating hormone, during menopausal transition. Due to limited follow up time, the age at the final menstrual period for many women in a study cohort is censored. A naive method that omits those subjects may yield biased estimation of the hormone profile from longitudinally observed hormone levels. We propose a two-stage pseudo likelihood approach to estimate the hormone profile during menopausal transition using a nonparametric stochastic mixed model. Simulations show that the proposed method corrects the bias and yields narrower confidence band.

Petkau, John (University of British Columbia)

Title: Neutralizing Antibodies and the Efficacy of Interferon Beta-1b in Multiple Sclerosis Clinical Trials

Abstract: Several large multi-center clinical trials have established that Type I interferons favorably influence clinical and MRI outcomes in relapsing-remitting multiple sclerosis (MS). However, some patients develop neutralizing antibodies (NAbs) to these treatments, reflecting an immune system response. The clinical significance of these NAbs has been unclear as titers vary widely and elevated NAb titers often diminish to undetectable levels. The question of whether NAbs impact on the efficacy of these treatments

is an unresolved scientific issue directly related to the question of how MS patients should be treated. It has also become a part of the marketing strategy of pharmaceutical companies with different Type I interferons approved for the treatment of MS: hundreds of millions of dollars per year are at stake.

In this presentation, I will discuss our involvement with this issue. I will describe the initial analyses which raised the concern, and the analyses we have carried out to try to resolve this issue. A fascinating part of our involvement has been attempting to persuade the neurological community of the need for more detailed analyses of the clinical trial data than is customary in the field to fully address this issue. Joint work with Rich White.

Qin, Li (Fred Hutchinson Cancer Center & University of Washington)

Title: A registration-based functional linear model for post-ART viral loads in patients with primary HIV infection

Abstract: In this paper, we introduce a registration based functional linear model to assess the post-Anti-retroviral therapy (ART) longitudinal viral loads of patients with primary HIV infections. Traditionally, such data were analyzed by approximate parametric or dynamical models. The parametric forms may be too restrictive while the dynamical models may be highly assumption dependent. The proposed model presents a trade-off between the parametric models and the flexible functional effects associated with the viral loads. In this framework, a test of the appropriateness of the parametric forms is also given. L-splines are used to model the viral loads and account for the plausible monotonicity in the curves over time. An O(N) algorithm is developed to fit the model efficiently. As a by-product, the derivative curves of the viral loads are obtained simultaneously which can be used to infer the rate of viral control and further verify the assumptions of the existing dynamical models. This is joint work with Wensheng Guo and Sarah Holte.

Qu, Annie (University of Illinois at Urbana-Champaign)

Title: Highly efficient aggregate unbiased estimating functions approach for correlated data with missing at random

Abstract: We develop a consistent and highly efficient marginal model for missing at random data using an estimating function approach. Our approach differs from inverse weighted estimating equations and the imputation method, in that our approach does not require estimating the probability of missing or imputing the missing response based on assumed models. The proposed method is based on an aggregate unbiased estimating function approach which does not require the likelihood function; however, it is equivalent to the score equation if the likelihood is known. The aggregate unbiased approach is based on a larger class of estimating functions than the pattern-unbiased approach. Therefore, the most efficient estimating function based on the aggregate unbiased approach is more efficient than any based on pattern-unbiased approaches. We provide comparisons of the three approaches using simulated data and also an HIV data example. Joint work with Bruce Lindsay and Lin Lu.

Rotnitzky, Andrea (Harvard University)

Title: Estimation and extrapolation of optimal dynamic treatment and testing strategies from observational longitudinal data.

Abstract: We review recent developments in the estimation of an optimal treatment strategy or regime from longitudinal data collected in an observational study. We propose novel methods for using the data obtained from an observational database in one health care system to determine the optimal treatment regime for biologically similar subjects in a second health care system when, for cultural, logistical, and financial reasons, the two health care systems differ (and will continue to differ) in the frequency of, and reasons for, both laboratory tests and physician visits. Finally, we propose a novel method for estimating

the optimal timing of expensive and/or painful diagnostic or prognostic tests. Diagnostic or prognostic tests are only useful in so far as they help a physician to determine the optimal dosing strategy, by providing information on both the current health state and the prognosis of a patient because, in contrast to drug therapies, these tests have no direct causal effect on disease progression. Our new method explicitly incorporates this no direct effect restriction. Joint work with James Robins and Liliana Orellana.

Song, Peter (University of Michigan)

Title: Analyzing Unequally Spaced Longitudinal Data with Quadratic Inference Functions

Abstract: Quadratic Inference Function (QIF) is getting increasingly popular, as an alternative to the well-known GEE method, to estimate parameters in the marginal models for longitudinal data. One limitation with the QIF is that it is currently only applicable for longitudinal data with equally spaced times. In this talk, a generalized QIF method is presented to relax this limitation. Both analytic and simulation results will be discussed for the proposed generalization.

Taylor, Jeremy (University of Michigan)

Title: Using joint models for longitudinal and survival data to give individual predictions

Abstract: In many medical studies there are two types of response variables, both of which may measure the progression of a disease or the response to an intervention. One type is a longitudinal variable and the other is an event time. In recent years there has been a growing literature on developing and using joint models to analyze such data. The longitudinal variable is typically modeled as a stochastic process or random effect model, and the event time is typically modeled as a time dependent proportional hazards model. There are many potential uses of joint models: to correct for measurement error in the covariate when considering the hazard model, to correct for bias in longitudinal modeling due to informative dropout, to use the longitudinal variable as an auxiliary or surrogate variable when the primary outcome is the event time. In this talk I will focus on using a joint model to assist with individual prediction of future event times for censored subjects. The model and methods are developed in the context of a prostate cancer application where the longitudinal variable is PSA and the event time is recurrence of the cancer following treatment with radiation therapy. Estimates of the parameters in the model are obtained by MCMC techniques. An efficient algorithm is developed to give individual predictions for subjects who were not part of the original data from which the model was developed. Many statistical issues arise including: how to validate the model, whether to use a parametric or a non-parametric longitudinal model, dependent censoring, what function of the longitudinal covariate path should be included in the hazard model, whether to provide measures of uncertainty associated with predicted probabilities.

Wang, Lu (University of Michigan)

Title: Nonparametric Regression in Longitudinal Studies with Dropout at Random

Abstract: We consider nonparametric regression for longitudinal data when some subjects drop out of the study at random. The reason people drop out may depend on the history of both outcome and covariates, but is independent of future outcome and covariates. We propose inverse probability weighted (IPW) kernel generalized estimating equations (GEEs) and IPW seemingly unrelated (SUR) kernel estimating equations using either complete cases or all available cases. None of these approaches require specification of a parametric model for the error distribution. We show that all these IPW kernel estimators are consistent when the probability of dropout is known by design or is estimated using a correctly specified parametric model. The most efficient IPW kernel GEE estimator is obtained by ignoring the within-subject correlation, while in contrast the most efficient IPW SUR kernel estimator is obtained by accounting for the within-subject correlation and is more efficient than the most efficient IPW kernel GEE

counterpart. The IPW kernel estimators obtained using all available cases are more efficient than those using complete cases when appropriate covariance matrices are used. We perform simulations to evaluate the finite sample performance of the proposed methods.

Wang, Naisyin (Texas A&M University and University of Michigan)

Title: Functional Latent Feature Regression Models for Data with Longitudinal Covariate Process

Abstract: We consider a joint model approach to study the association of nonparametric latent features of multiple longitudinal processes with a primary endpoint. We propose estimation procedures and corresponding supportive theory that allow one to perform investigation without making distributional assumption of the latent features. The uncertainty which is associated with accounting for these latent features is also properly taken into consideration. We investigate the practical implications behind certain theoretical assumptions, which aims at having a better understanding of where the estimation variation lies. Numerical performances of the proposed approach were illustrated through simulations and a hypertension study.

Zhao, Yang (University of Regina)

Title: Likelihood Method for Regression Models with Data Missing at Random

Abstract: For regression models with covariate data missing at random the maximum likelihood (ML) methods proposed by Ibrahim (1990), Chen (2004), and Zhang and Rockette (2005) do not require modeling the missing data probability and parametric assumptions of the covariates. The result estimators are therefore totally robust. In this talk we consider extending the ML method to deal with missing data problems in longitudinal data analysis.

PARTICIPANTS
(in alphabetic order by last names)

Participant	Affiliation	E-mail
Abarin, Taraneh	Samuel Lunenfeld Research Institute, Toronto	abarin@lunenfeld.ca
Carpenter, James	London School of Hygiene & Tropical Medicine	James.Carpenter@lshtm.ac.uk
Carroll, Raymond	Texas A&M University	carroll@stat.tamu.edu
Chen, Baojiang	University of Washington	baojiangchen@gmail.com
Chen, Zhijian	University of Waterloo	z11chen@math.uwaterloo.ca
Cook, Richard	University of Waterloo	rjcook@uwaterloo.ca
Dean, Charmaine	Simon Fraser University	dean@stat.sfu.ca
He, Wenqing	University of Western Ontario	whe@stats.uwo.ca
Hu, Joan	Simon Fraser University	joanh@sfu.ca
Ibrahim, Joseph	University of North Carolina	ibrahim@bios.unc.edu
Kenward, Mike	London School of Hygiene & Tropical Medicine	mike.kenward@lshtm.ac.uk
Lai, Tze Leung	Stanford University	lait@stanford.edu
Lee, Youngjo	Seoul National University	youngjo@snu.ac.kr
Li, Daniel	University of Manitoba	umlih2@cc.umanitoba.ca
Li, Song (Sebastian)	University of Toronto	Sebastian.Li@utoronto.ca
Liang, Hua	University of Rochester	Hua_Liang@URMC.Rochester.edu
Lin, Xihong	Harvard University	xlin@hsph.harvard.edu
Little, Rod	University of Michigan	rlittle@umich.edu
Liu, Wei	York University	liuwei@mathstat.yorku.ca
Liu, Juxin	University of Saskatchewan	liu@math.usask.ca
Ma, Renjun	University of New Brunswick	renjun@unb.ca
McCulloch, Charles	University of California, San Francisco	CMcCulloch@epi-ucsf.org
Nan, Bin	University of Michigan	bnan@umich.edu
Nathoo, Farouk	University of Victoria	nathoo@math.uvic.ca
Neuhaus, John	University of California, San Francisco	john@biostat.ucsf.edu
Ory, Marcia	Texas A&M Health Science Center	mory@srph.tamhsc.edu
Petkau, A. John	University of British Columbia	john@stat.ubc.ca
Qin, Li	Fred Hutchinson Cancer Research Center	lqin@scharp.org
Qiu, Jin	Zhejiang University of Economics and Finance	qiujin_71@hotmail.com
Qu, Annie	University of Illinois, Urbana Champaign	anniequ@illinois.edu
Rotnitzky, Andrea	Harvard University & Di Tella University	arotnitzky@utdt.edu
Song, Peter	University of Michigan	pxsong@umich.edu
Sun, Jiayang	Case Western Reserve University	jiayang@sun.STAT.cwru.edu
Taylor, Jeremy	University of Michigan	jmgt@umich.edu
Wang, Liqun	University of Manitoba	wangl1@cc.umanitoba.ca
Wang, Naisyin	Texas A&M University	nwangstat@gmail.com
Wang, Lu	University of Michigan	luwang@umich.edu
Wu, Lang	University of British Columbia	lang@stat.ubc.ca
Xu, Ronghui	University of California, San Diego	rxu@math.ucsd.edu
Yi, Grace	University of Waterloo	yyi@math.uwaterloo.ca
Zeng, Leilei	Simon Fraser University	lzeng@sfu.ca
Zhao, Yang	University of Regina	Yang.Zhao@uregina.ca

DIRECTIONS

The Banff International Research Station (BIRS, *http://www.birs.ca*) is located in *Corbett Hall* at the Banff Centre in the town of Banff, Alberta, Canada. A wealth of information about the centre, the town and the surroundings can be found at *http://www.banffcentre.ca/Conferences/visitor*, which also provides detailed directions and maps.

Banff is nestled in the Rocky Mountains, approximately 1.5 hours west of Calgary and can be reached by car or by transit from the Calgary International Airport.

Banff/Calgary Airport Shuttles

Most workshop participants will be arriving and departing from the *Calgary International Airport.* There are direct shuttles operating between the Calgary International Airport and Banff (all include a stop at the Banff Centre PDC). *Reservations are recommended (and are necessary for the trip returning to the airport).*

The Banff-Calgary Airporter Shuttle: Call 1-888-HIWAY-01 (1-888-449-2901) in the US and Canada, or visit the website at *http://www.banffairporter.com* for reservations, schedules, and fares. BIRS participants are entitled to a 10% discount. To obtain the discount when booking on the web, use

```
        User name: birs               Password:  birs04
```

and follow the instructions.

Directions by Car

If you are renting a car then you should be aware that in order to enter Banff National Park (in which BIRS is located) you need to pay a daily user fee. For more on this you may go to the Parks Canada website: *http://www.birs.ca/participants/transportation.* You will pay this as you enter into the National Park. Visitors who come by the shuttle bus are not subject to Park fees. The Banff Centre will provide free parking for you for your stay at BIRS. For driving direction, see *http://www.birs.ca/participants/transportation.*

Check-in

To check-in and obtain your room key, upon arrival proceed to the registration desk in the Professional Development Centre. This is where the shuttle service will drop you off. The registration desk is open 24 hours so you may check-in anytime. Registration should be straight-forward as the Banff Centre should be expecting your arrival. They will ask you for a **credit card** imprint. This is only to cover your incidental expenses, and to allow you to use the telephone in your bedroom. Without it your phone will not work.

After you have registered and received your key, please proceed to **Corbett Hall**, which is just to the right and back of the building on which you can see flags of various countries. Corbett Hall is the home of BIRS and there you will be able to find more information about the many facilities available to you at the Banff Centre. *Every bedroom at BIRS has a telephone and a computer terminal with high-speed access to the Internet. There are wireless access points in some common areas.*

Meeting Facilities

Meeting facilities are in the Max Bell Building which adjoins Corbett Hall. For more information about BIRS and its facilities, please visit: *http://www.birs.ca/facilities/facilities.*

Maps

Maps for the town of Banff and for the Banff Centre are available at *http://www.birs.ca/participants/maps.*