KATIA CHARLAND, McGill University and Children's Hospital Boston
An application of gravity models to identify factors related to community pandemic influenza A/H1N1 vaccine coverage

Nineteen vaccination clinics were established in Montreal, Canada, as part of the 2009 A/H1N1p mass vaccination campaign. Though approximately 50 percent of the population was vaccinated (e.g. compared to 8 percent in France), there was considerable geographic variation in vaccine coverage. Analysis of the geographic variation in healthcare utilization could potentially reveal underlying barriers of access to healthcare services. In this talk I discuss an application of gravity models to identify characteristics of the communities and the mass vaccination clinics that were associated with vaccine uptake. Gravity models are well suited to this problem as they examine the features of origin and destination that affect traffic, or flow, from origin to destination. Typically, the shorter the distance between origin and destination and the larger the mass of the origin/destination (e.g. population size/clinic capacity), the greater the 'gravitational pull' or flow between origin and destination. We identified several factors associated with rates of vaccinations. For example, communities in which only a small proportion of the population spoke English or French, tended to have low vaccine coverage. Clinics placed in materially deprived neighbourhoods with high residential density and high violent crime rates did not perform well, in general.

JAMES DAI, Fred Hutchinson Cancer Research Center
Adherence, ARV drug concentration and estimation of PrEP efficacy for HIV prevention

Assays to detect antiretroviral drug levels in study participants are increasingly popular in PrEP trials as they provide an objective measure of adherence. Current correlation analyses of drug concentration data are prone to bias because the comparisons are not protected by randomization. In this talk, I will discuss the causal estimand of prevention efficacy among drug compliers, those who would have had a level of drug concentration had they been assigned to the drug arm. Both dichotomous drug detection status and continuous drug concentration measure are considered. The identifiability of the causal estimand is facilitated by either exploiting the exclusion restriction that drug noncompliers do not acquire any prevention benefit, or imputing drug measure by correlates of adherence. For the former approach, we develop sensitivity analysis that relaxes the exclusion restriction. For the latter approach, we study the performance of regression calibration. Applications to published data from existing PrEP trials suggest high efficacy estimates among drug compliers. In summary, the proposed inferential method provides an unbiased assessment of PrEP efficacy among drug compliers, thus adding to the primary intent-to-treat analysis.

RAPHAEL FONTENEAU, University of Liège (Belgium) / Inria Lille - Nord Europe (France)
Batch Mode Reinforcement Learning based on the Synthesis of Artificial Trajectories

Batch mode reinforcement learning (BMRL) is a field of research which focuses on the inference of high-performance control policies when the only information on the control problem is gathered in a set of trajectories. Such situations occur for instance in the case of clinical trials, for which data are collected in the form of batch time series of clinical indicators. When the (state, decision) spaces are large or continuous, most of the techniques proposed in the literature for solving BMRL problems combine value or policy iteration schemes from the Dynamic Programming (DP) theory with function approximators representing (state-action) value functions. While successful in many studies, the use of function approximators for solving BMRL problems has also drawbacks. In particular, the use of function approximator makes performance guarantees difficult to obtain, and does not systematically take advantage of optimal trajectories. In this talk, I will present a new line of research for solving BMRL problems based on the synthesis of "artificial trajectories" which opens avenues for designing new BMRL algorithms. In particular, it avoids the two above-mentioned drawbacks of the use of function approximator.

BRIAN INGALLS, University of Waterloo
Sensitivity Trade-offs in Systems Biology

The stabilizing effect of negative feedback is key to biological self-regulation (homeostasis). Feedback allows a system to maintain its preferred behaviour in an unpredictable environment. The prevailing wisdom is that negative feedback typically stabilizes a system (making it less sensitive to external perturbations), while positive feedback is destabilizing (i.e. it increases sensitivity). Of course, negative feedback can also generate instability, for example in producing oscillations. However, even when acting to improve a system's robustness, negative feedback typically redistributes sensitivity within a network, rather than directly reducing it. In some cases, this redistribution is governed by an explicit constraint: a conservation of sensitivity. This talk will introduce sensitivity conservation statements commonly used in control engineering and molecular systems biology, and introduce a unifying formulation.

MICHAEL MACKEY, McGill University
Molecular distributions in gene regulatory networks

Extending the work of Friedman et al. (2006), we study the stationary density of the distribution of molecular constituents in the presence of noise arising from either bursting transcription or translation, or noise in degradation rates. We examine both the global stability of the stationary density as well as its bifurcation structure. We have compared our results with an analysis of the same model systems (either inducible or repressible operons) in the absence of any stochastic effects, and shown the correspondence between behaviour in the deterministic system and the stochastic analogs. We have identified key dimensionless parameters that control the appearance of one or two stable steady states in the deterministic case, or unimodal and bimodal densities in the stochastic systems, and detailed the analytic requirements for the occurrence of different behaviours. This approach provides, in some situations, an alternative to computationally intensive stochastic simulations. Our results indicate that, within the context of the simple models we have examined, bursting and degradation noise cannot be distinguished analytically when present alone.

DAVID MCMILLEN, University of Toronto Mississauga
Design methods and constraints for biological integral control

Synthetic biology includes an effort to use design-based approaches to create novel "controllers," biological systems aimed at regulating the output of other biological processes. The design of such controllers can be guided by results from control theory, including the strategy of integral feedback control, which is central to regulation, sensory adaptation, and long-term robustness. Realization of integral control in a synthetic network is an attractive prospect, but the nature of biochemical networks can make the implementation of even basic control structures challenging. Here we present a study of the general challenges and important constraints that will arise in efforts to engineer biological integral feedback controllers, or to analyze existing natural systems. Constraints arise from the need to identify target output values that the combined process plus controller system can reach; and to ensure that the controller implements a good approximation of integral feedback control. These constraints depend on mild assumptions about the shape of input-output relationships in the biological components, and thus will apply to a variety of biochemical systems. We summarize our results as a set of variable constraints intended to provide guidance for the design or analysis of a working biological integral feedback controller.

THEODORE PERKINS, Ottawa Hospital Research Institute
What Do Molecules Do When We're Not Looking? State Sequence Analysis for Stochastic Chemical Systems

Many biomolecular systems depend on orderly sequences of chemical transformations or reactions to carry out their functions. Yet, the dynamics of single molecules or small-copy-number molecular systems are significantly stochastic. I will describe State Sequence Analysis, a new approach for predicting or visualizing the behaviour of stochastic molecular systems by computing maximum probability state sequences based on initial conditions or boundary conditions. I demonstrate this approach by analyzing the acquisition of drug-resistance mutations in the HIV genome, which depends on rare events occurring on the

timescale of years, and the stochastic opening and closing behaviour of a single sodium ion channel, which occurs on the timescale of milliseconds. In both cases, the approach yields novel insights into the stochastic dynamical behaviour of these systems, including insights that are not correctly reproduced in standard time-discretization or stochastic-simulation approaches to trajectory analysis.

JANET RABOUD, University of Toronto
Left truncation in the context of competing risks

An analysis involving left truncation in the context of competing risks will be presented. The goal of the analysis is to estimate the effect of co-infection with hepatitis C on the risk of developing cardiovascular (CVD) disease among HIV infected individuals. Time is measured from the date of initiation of antiretroviral therapy. Non CVD deaths are a competing risk to the event of interest. Left truncation occurs due to the fact that antiretroviral therapy may have been initiated before enrolment into the cohort. While left truncation of the competing risk, non-CVD death, is complete, some information is available on CVD events prior to enrolment into the study through the collection of medical histories. The degree of completeness of this data varies by calendar time and site, in this multi-site cohort study. The analysis is further complicated by the desire to model hepatitis C with a time dependent variable, since infection may clear spontaneously or with treatment. Results of this work-in-progress analysis will be presented, as well as plans for further work.

MARC ROUSSEL, University of Lethbridge
Stochastic effects in gene transcription

Gene transcription is typically the major source of noise in gene expression. We have developed models of gene transcription for both prokaryotes and eukaryotes. These models allow us to examine the effects of the kinetics of various elementary reaction steps on the overall statistical behavior of transcription, and in particular on the distribution of transcription times. Here we review a few results obtained from our models, emphasizing how these results impact large-scale gene network modeling.

ANDREW RUTENBERG, Dalhousie University
Stochastic Models of Plastic Development in Cyanobacterial Filaments

When deprived of fixed nitrogen, filamentous cyanobacteria differentiate nitrogen fixing heterocyst cells in a regular pattern. By including uniform cellular fixed-nitrogen storage in a filamentous model of nitrogen dynamics, growth, and heterocyst differentiation we can explain the stochastic timing of heterocyst commitment. Stochasticity arises mostly from the natural population structure of cell lengths in the filament. Later events in heterocyst differentiation were consistent with deterministic heterocyst development following commitment. Our computational model has qualitatively reproduced many of the measurements associated with heterocyst differentiation including the initial and steady state heterocyst patterns.

MARK VAN DER LAAN, UC Berkeley
Statistical Methods for Causal Inference In HIV Research

In this talk I will review some of the statistical challenges we encountered in our collaborations in HIV research. One collaboration was concerned with the assessment of the effects of mutations in the HIV virus on drug-resistance, involving interval censored time to event outcomes, confounding by the patient history and other mutations. In another collaboration we are concerned with estimation of subgroup-specific causal effects of treatment on time to death and viral failure based on a randomized controlled trial in which subjects are lost to follow up in response to time-dependent markers. We have also worked on estimation of individualized rules for when to switch a drug regimen and when to start treatment for HIV infected patients based on a variety of observational studies. Currently, we are involved in designing a RCT for comparing a "treat early" intervention with the current standard w.r.t. HIV prevention at the community level, and in determining optimal rules for triggering HIV testing based on observing the history of subjects including their adherence profile. We will demonstrate that we employed a general roadmap for targeted learning of causal effects, involving the most recent advances in modeling, estimation, and inference.

YONGLING XIAO, Institut national d'excellence en santé et en services sociaux (INESSS)
Flexible marginal structural models for estimating cumulative effect of time-varying treatment on the hazard

Many longitudinal studies deal with both time-dependent (TD) exposures or treatments and TD confounders. When a TD confounder acts also as a mediating variable, marginal structural Cox models (Cox MSM) can be used to consistently estimate the causal effect of the TD exposure. On the other hand, modeling of the effect of a TD exposure requires specifying the relationship between the hazard at time t and the entire past exposure history, up to time t. Flexible modeling of the weighted cumulative exposure (WCE) has been proposed to address this challenge. However, the existing WCE models do not permit accurate adjustment for TD confounding/mediating variables, while the existing MSMs do not incorporate flexible estimation of the cumulative effects of TD exposures.

In this study, we propose a flexible marginal structural Cox model with weighted cumulative exposure modeling (WCE MSM), which combines the Cox MSM and WCE approaches, thereby simultaneously addressing the two aforementioned analytical challenges. Specifically, by controlling for confounders using the inverse-probability-of-treatment weights and estimating the WCE with cubic regression splines, the new WCE MSM can estimate the total causal treatment effect, that accounts for both direct cumulative effects of past treatments and their 'indirect effects', mediated by the TD mediators.

Simulation results confirm that the proposed WCE MSM yields accurate estimates of the causal treatment effect under settings of complex exposure effects and time-varying confounding. The new method was applied to the Swiss HIV Cohort Study data to reassess the association between antiretroviral treatment abacavir (ABC) and cardiovascular risk.

JESSICA YOUNG, Harvard School of Public Health
Simulation from a known Cox MSM using standard parametric models for the g-formula

Inverse probability (IP) weighted estimation of Cox Marginal Structural Models (MSMs) is now a popular approach for estimating the effects of time-varying antiretroviral therapy regimes on survival in HIV-infected patients. Unlike standard estimates, IP weighted estimates of the parameters of a correctly specified Cox MSM may remain unbiased for the causal effect of following one regime over another in the presence of a time-varying confounder affected by prior treatment (e.g. CD4 cell count). A standard estimate might be a likelihood-based estimate of the parameters on treatment in a time-dependent Cox model for the observed failure hazard at each time conditional on past measured treatment and confounders. Previously proposed methods for simulating data according to a known Cox MSM are useful for studying the performance of an IPW estimator as they involve explicit knowledge of quantities required for unbiased IPW estimation. These approaches are limited, however, for studying bias in a standard estimate due only to the presence of time-varying confounders affected by prior treatment as they lack explicit knowledge of the observed conditional failure hazard. Here an alternative approach to Cox MSM data generation is considered that addresses this problem by generating data from a standard parametrization of the observed data distribution. In this case, the true Cox MSM parameters may be derived by the relation between a Cox MSM and the g-formula. This talk will review this relationship in general and work through an example. This approach has limitations including those implied by the g-null paradox theorem.

DANIEL ZENKLUSEN, Université de Montréal
Imaging Single Transcripts Resolves Kinetics of Gene Expression Processes

Many cellular processes involve a small number of molecules and undergo stochastic fluctuations in their levels of activity; consequentially, biological processes are probably not executed with the precision often assumed. Hence, expression levels of proteins and mRNAs can vary considerably over time and between individual cells, limiting considerably the value of experimental datasets acquired through ensemble measurements that rely in pooling thousands of cells, as these datasets will only reflect an average behavior of a particular process. This underlines that cells should not be studied as 'the average cell' but as individual entities. Truly understanding 'cellular biochemistry' therefore requires the ability to study the behavior of individual cells, and ideally, individual molecules, only this results in datasets that represent the whole range of possible scenarios that can occur within a cell. We will summarize recent advances in single molecule RNA imaging approaches that have facilitated single molecule studies in cells and illustrate how we apply these techniques to determine how cells execute different processes along

the gene expression pathway, in particular transcription.