Information and Knowledge for Decision Making

An NSF I/UCRC Planning Grant Workshop

I/UCRC
INDUSTRY & UNIVERSITY COOPERATIVE RESEARCH PROGRAM

Research Concept Presentations

Section 1

VDI Center for Visual & Decision Informatics

renci

NSF

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

L.I.F.E. Form Access

Level of Interest and Feedback Evaluation (LIFE) Forms

- Please go to: http://iucrc.renci.org
- Select Planning Grant Workshop
- Then select L.I.F.E Form Evaluations
    - PASSWORD: **unc2015**
    - **ID yourself as IAB**

Multi-Agent Systems for Collective Intelligence & Decision-Making

Noel P. Greis

Kenan-Flagler Business School

ISUCRC
INDUSTRY & UNIVERSITY COOPERATIVE RESEARCH PROGRAM

OBJECTIVES

To solve problems that are complex, dynamic, and involve multiple actors each with awareness of and control over their environment, each with separate ,and often competing, goals. Typical problems that can be addressed by MAS include:

- Competitive dynamics of an industry;
- Complex market behaviors (c.f. healthcare, energy, etc.);
- Allocation of resources across a large enterprise (c.f. supply chain management);
- Complex scheduling (c.f. emergency department, production lines)
- Operation of complex physical systems (c.f. smart grids);
- Analysis of operational efficiencies (c.f. healthcare delivery)

APPROACH/TECHNIQUES

Combine agent-based technology with other tools such as logic programming and machine learning in a decision-theoretic framework:

- Multi-Agent Systems offers a methodology for representing and modeling complex systems of actors or entities;
- Logic Programming describes the behavior of each actor/agent which are assumed to behave “rationally” or in some way that can be “expected” or “described”; and
- Machine Learning for capturing the experiential knowledge of the individual agents acting within the dynamic system.

DELIVERABLES

- Simulation tools customized for specific business applications that provides critical input for business decisions;
- Visualization tools for understanding the behaviors of complex distributed systems; and
- Recommendations for providing targeted & actionable intelligence.

BENEFITS TO INDUSTRY

- Offers an approach to model complex systems;
- Structure of MAS resembles organization or enterprise or organizational system;
- Reflects decentralized problem-solving within an organization;
- Efficiently retrieves, filters, and globally coordinates information from sources that are organizationally & spatially distributed.

Multi-Agent Systems for Collective Intelligence & Decision-Making

Noel P. Greis

Kenan-Flagler Business School

I/UCRC
INDUSTRY & UNIVERSITY COOPERATIVE RESEARCH PROGRAM

SIMULATION METHODS FOR DECISION-MAKING FOR HEALTHCARE SYSTEMS

TOOLKIT

System Dynamics Modeling
High-level aggregation of the objects being modeled
Systems-oriented with no representation of individual "entities".

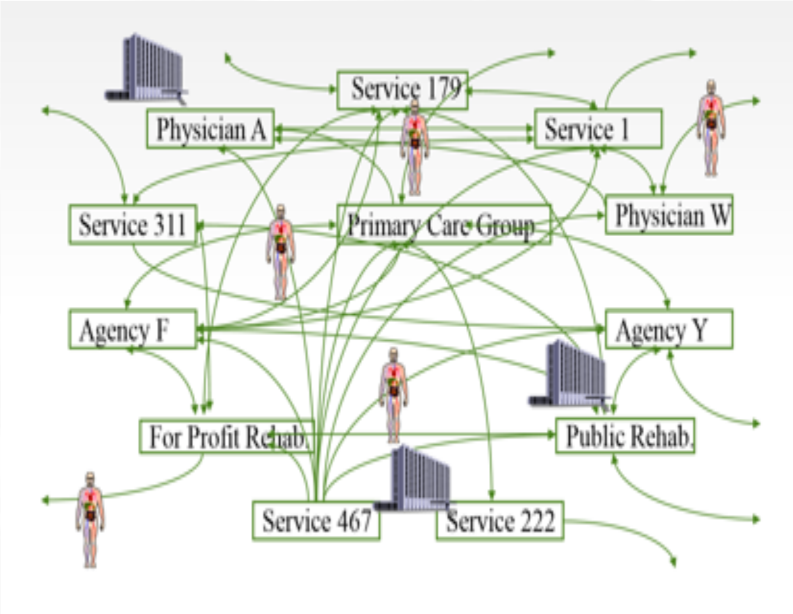
Discrete Event Modeling
Continuous real-world processes
Process-oriented where entities are the same with no individual behaviors.

AGENT-BASED MODELING
Decentralized, Individual-centric
Focus is on modeling individual entities and the interactions between them.

HEALTHCARE AS A COMPLEX ADAPTIVE SYSTEM

A system composed of many diverse and autonomous entities or agents which are interrelated, interdependent, and linked … and that learn from experience and adjust (not just react) to changes in their environment.

- **Behavior at the "Agent" level**
  - ❑ Numerous
  - ❑ Heterogeneous
  - ❑ Local Interactions
  - ❑ Nestedness
  - ❑ Adaptiveness
- **Behavior at the "System" level**
  - ❑ Emergence
  - ❑ Self-Organization
  - ❑ Co-Evolution
  - ❑ Path Dependency



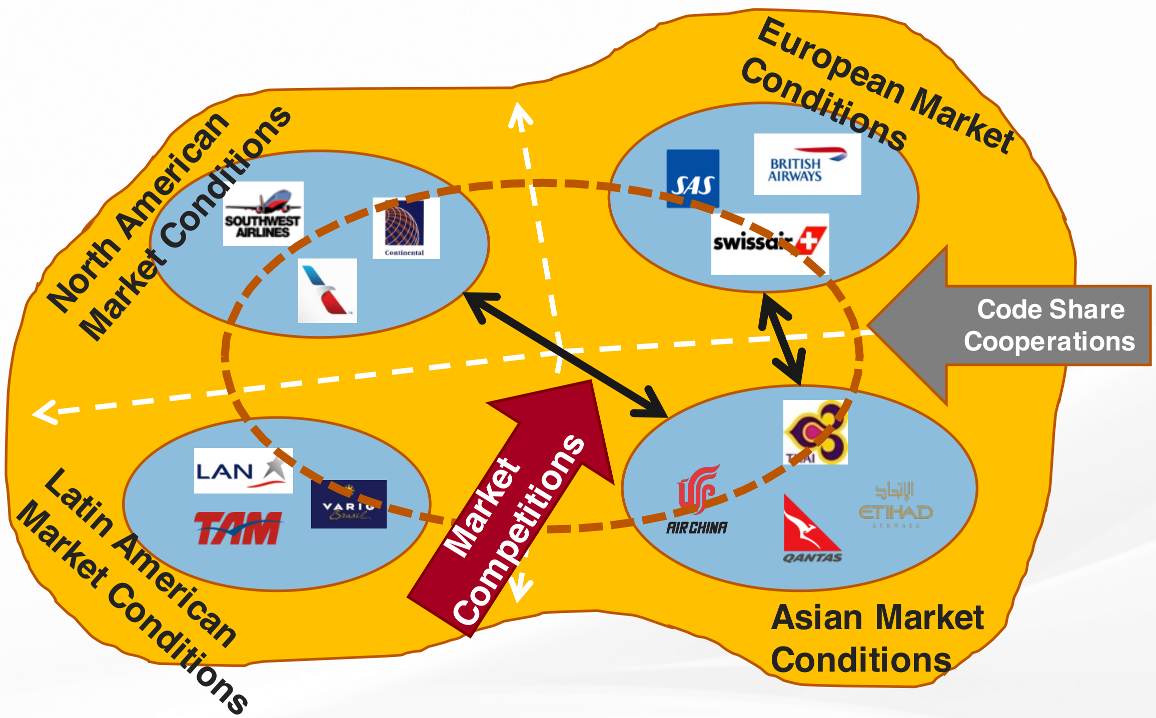
Multi-Agent Systems for Collective Intelligence & Decision-Making

Noel P. Greis
Kenan-Flagler Business School

I/UCRC
INDUSTRY & UNIVERSITY COOPERATIVE RESEARCH PROGRAM

COMPETITIVE DYNAMICS

Modeling Organizational Competition and Collaboration Using Agent-Based Models



CENTRAL PREMISE

An organization's future behavior can be understood within the context of its *previous decisions and relevant business context*. Agent-based models that capture these contextual behaviors can inform *predictions of future behaviors* as market and economic conditions change.

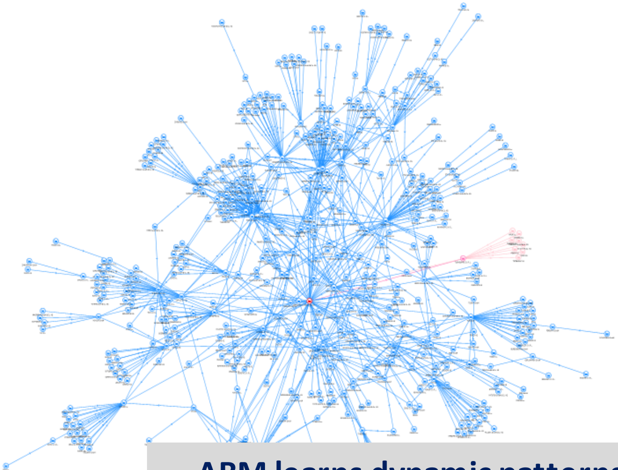
- Autonomous ***software agents*** represent airlines (or other companies)
- Individual airline agent behavior is modeled using Logic Programming based on ***learned previous history.***
- Airline agents engage in ***competitive and collaborative*** behaviors.
- Airline agents respond to ***changes in market environment*** (c.f. jet fuel price, GDP, passenger demand, etc.)

GOAL OF ABM
ANTICIPATE INDUSTRY AND COMPETITOR FUTURE BEHAVIOR FOR BETTER DECISION-MAKING

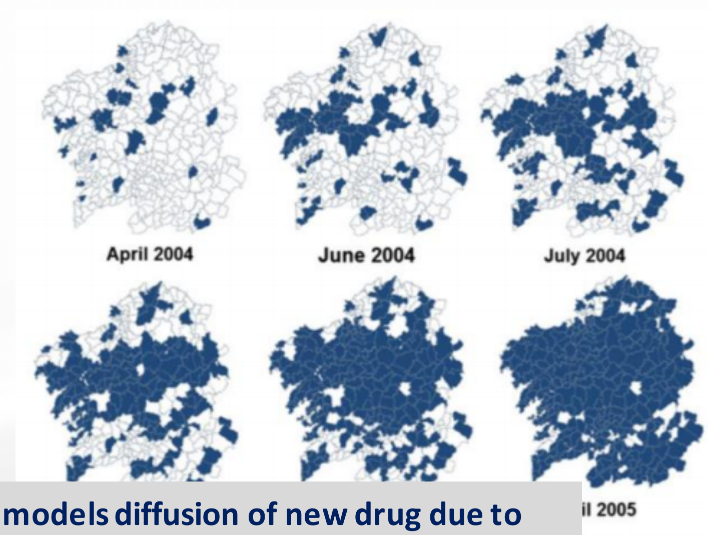
Multi-Agent Systems for Collective Intelligence & Decision-Making

Noel P. Greis

Kenan-Flagler Business School



ABM learns dynamic patterns of supply and demand to "optimize" their performance to avoid stock-outs....Greis et al. in progress.



ABM models diffusion of new drug due to imitative behavior and social interaction among prescribers...Pombo-Ribero et al.

EXAMPLE ABM APPLICATIONS IN HEALTHCARE

- Dynamics of healthcare delivery interventions
- Uptake of new or alternative health policies
- Introduction of new drug in a market
- Behavior of physician teams and impact on patient outcomes
- Flows of health information on social networks

Healthcare improvement efforts often focus on changing the behavior of individuals. ABM changes the focus from the individual to the RELATIONSHIPS and INTERDEPENDENCIES among individuals in the system.

Interactive Visual Cohort Selection Methods for Large Scale Temporal Event Data

David Gotz

School of Information and Library Science
Carolina Health Informatics Program

OBJECTIVES

- Develop novel visual analysis methods designed to support high-speed, intuitive, and statistically rigorous studies using large-scale and high-dimensional temporal event datasets. The developed methods will support event data alignment and cohort selection, with a focus on challenges related to high-dimensionality and data variety (sparse, irregular over time, etc.).

APPROACH/TECHNIQUES

- Scalable visual analytics methods will be developed to support alignment of temporal event records, and, using the aligned data, exploratory high-dimensional visual cohort selection. These methods will combine interactive data visualization methods, statistical analysis, and human-computer interaction techniques to make tasks intuitive, high-speed, and robust.

DELIVERABLES

- New visual analysis methods (including both design and algorithmic foundations) would be documented and described in formal reports. This may take the form, for example, of a peer-reviewed publication.
- In addition, prototype software will be made available to members for use as the basis for new products or for integration with other existing technologies.

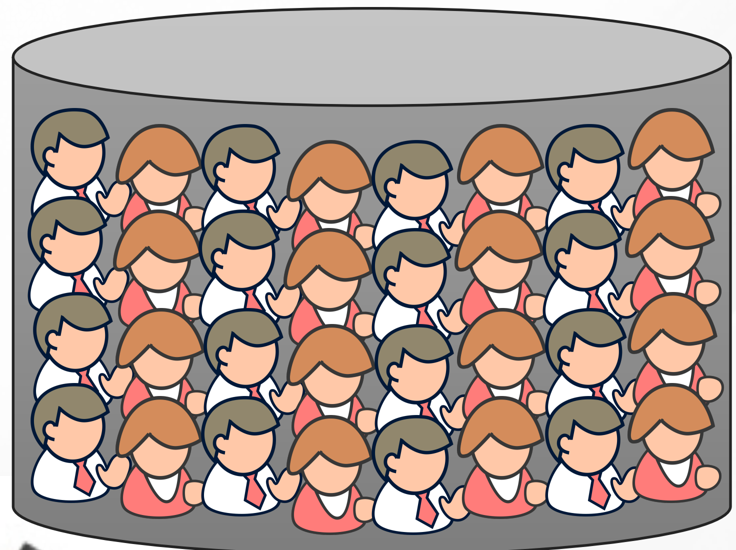
BENEFITS TO INDUSTRY

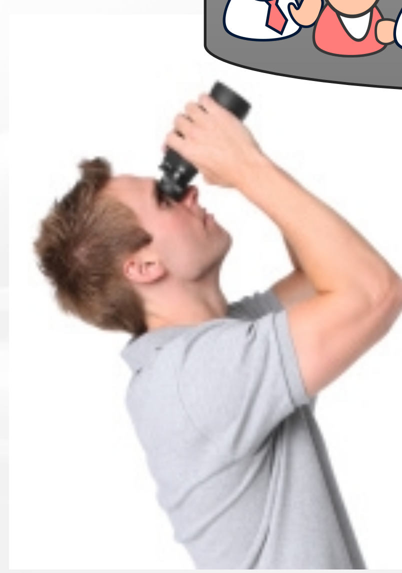
- New techniques and software tools would be available for data-driven analysis and decision making. These tools would have applications in domains such as (1) social media analysis / customer management, (2) systems log file analysis, and (3) security applications.
- Specific to healthcare, these tools enable a rich set of population health or precision medicine applications which are emerging as critical capabilities in the Learning Health System era.

David Gotz

School of Information and Library Science
Carolina Health Informatics Program

Visual Methods for Temporal Event Analysis





- Millions of patients
- 100,000+ dimensions
- Sparsely populated
- Irregularly sampled over time
- "Real world" data
    - ***Unknown contents***
    - ***Sampling biases***

Interactive Visual Cohort Selection Methods for Large Scale Temporal Event Data

David Gotz

School of Information and Library Science
Carolina Health Informatics Program

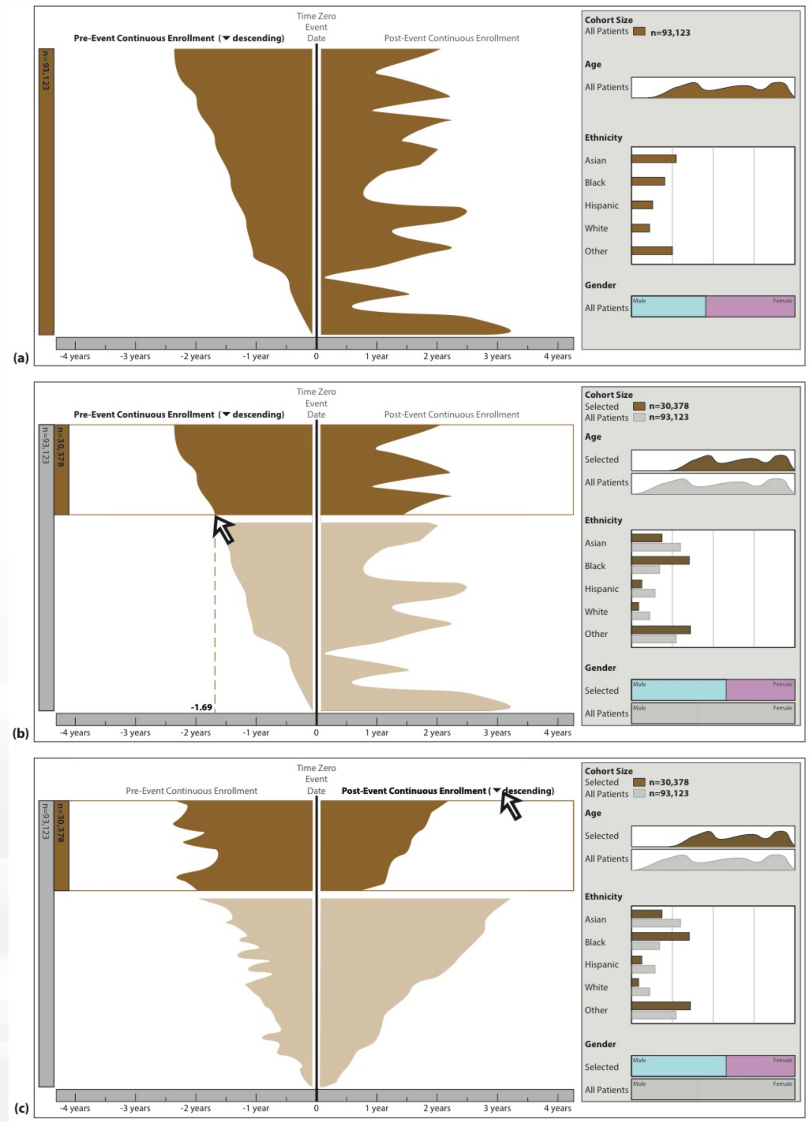
I/UCRC
INDUSTRY & UNIVERSITY COOPERATIVE RESEARCH PROGRAM

Aim 1: Visual Analytics for Alignment and Integration

- Temporal data must be aligned
  - Sentinel events: individual or patterns
  - Define "Time Zero"

- Time window must be defined for both pre- and post- alignment periods
  - Real-world data is highly variable
  - Tradeoff between:
    1. **Window size**
    2. **Cohort size**
    3. **Makeup of cohort**

- Must understanding confounds (e.g., impact of alignment/windowing on demographics)



Design Sketch

Interactive Visual Cohort Selection Methods for Large Scale Temporal Event Data

David Gotz
School of Information and Library Science
Carolina Health Informatics Program

Aim 2: Cohort Selection for High-Dimensional Data

- Complex sets of inclusion / exclusion criteria
    - High-dimensional variable space
    - Constructed features
        - Aggregation (e.g., hierarchy)
        - Temporal patterns
- Exploratory, iterative process
    - Each constraint narrows the data set
    - Dependencies between variables produces changes in variable distribution
    - Methods needed to assess—and correct for—changes in representation

Proof-of-concept cohort selection system
(Nov 2015)

Maintaining Exercise Through the Cardiac Rehabilitation Transition

Kelly R. Evenson

Gillings School of Global Public Health
Department of Epidemiology

INDUSTRY & UNIVERSITY COOPERATIVE RESEARCH PROGRAM

OBJECTIVES

- Develop and test a theory-driven intervention designed to improve exercise adherence during and following completion of a supervised cardiac rehabilitation program. This intervention would be designed to be easily scaled up and disseminated widely.

APPROACH/TECHNIQUES

1. Interviews with cardiac rehabilitation program directors and cardiac rehabilitation patients.
2. Identification of the most promising tools to use in cooperation with activity tracking that addresses the needs of the cardiac rehabilitation patient.
3. Collect pilot data on cardiac rehabilitation patients and develop ways of analyzing their exercise patterns.
4. Develop options for providing feedback to patients based on their exercise patterns.

DELIVERABLES

1. Develop procedures to work with cardiac rehabilitation centers to integrate the intervention into their existing program
2. Develop protocols to analyze activity tracking data and provide feedback to the patient both during and following cardiac rehabilitation.

BENEFITS TO INDUSTRY

1. For insurance companies that pay for cardiac rehabilitation, a successful intervention would help the patient maintain the health benefits gained from cardiac rehabilitation, and thus save in claims costs.
2. The intervention approach to assist patients in successfully transitioning from a structured rehabilitation program to their own could be applied to other types of rehabilitation in order to save in claims costs.

Maintaining Exercise Through the Cardiac Rehabilitation Transition

Kelly R. Evenson
Gillings School of Global Public Health
Department of Epidemiology

ISUCRC
INDUSTRY & UNIVERSITY COOPERATIVE RESEARCH PROGRAM

Cardiac rehabilitation patients:

CABG, angioplasty/stent in coronary artery, MI (heart attack), stable angina, heart/lung transplant, and heart valve replacement/repair. May also include congestive heart failure, following pacemaker or automatic internal cardiac defibrillators (AICD) insertion, multiple CVD risk factors, etc.

Phase I ----→ Phase II ---→ Phase III or on own

Kelly R. Evenson

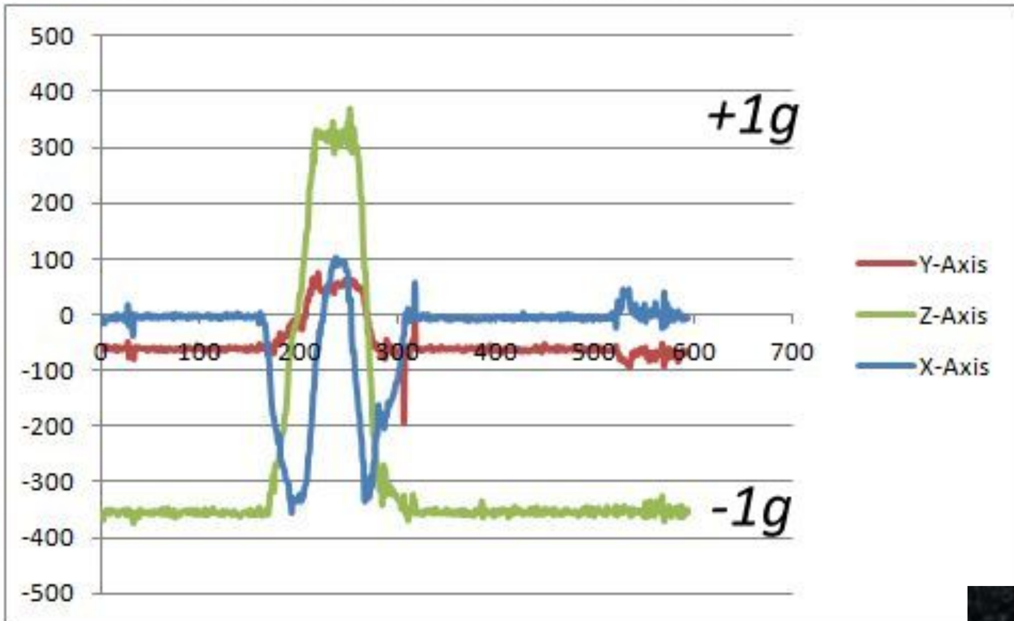
Gillings School of Global Public Health
Department of Epidemiology

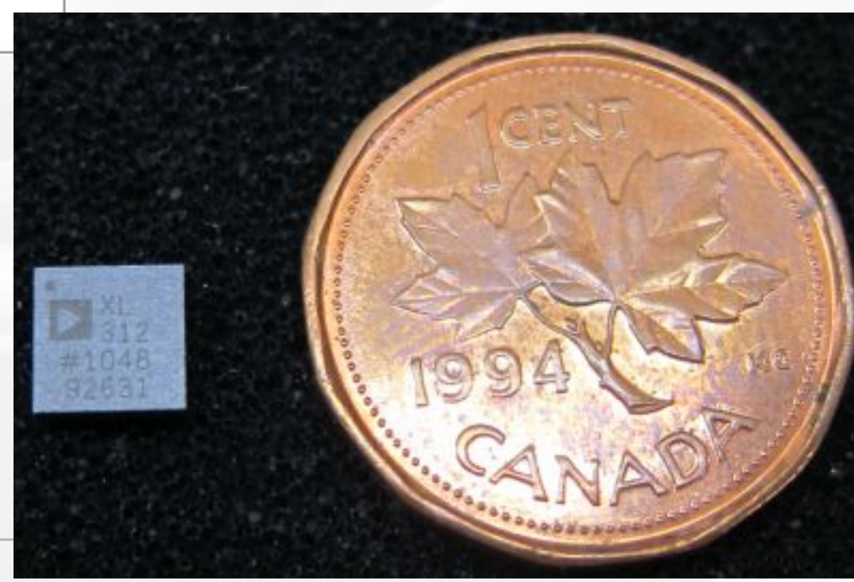
Activity/Fitness Trackers: Consumer wearable devices used for monitoring fitness- and physical activity-related measures that provide feedback via a mobile device, base station, or computer for long-term tracking and data storage

Kelly R. Evenson

Gillings School of Global Public Health
Department of Epidemiology

- Accelerometry





Prediction of Complex Genetic Traits

Kirk Wilhelmsen

School of Medicine - Genetics

I/UCRC INDUSTRY & UNIVERSITY COOPERATIVE RESEARCH PROGRAM

OBJECTIVES

- Prediction of Complex (non-mendelian) traits
  - Height
  - Body Mass
  - Which drug should a patient use ?
  - When will a patient in a primary prevention trial for Alzheimer's Disease be expected to develop disease?

APPROACH/TECHNIQUES

- Data input
  - Measured or imputed sequence of DNA at 8-10 million basepairs that are commonly variable and actual human phenotypes for thousands of patients.
  - Simulated sequence and realistic modeled phenotypes for unlimited numbers of subjects
- Genomic prediction using many methods

DELIVERABLES

- Head to comparison of genomic prediction
- Data simulator that is of general interest for problems where N<<P
- Integration of genomic data from real data that has been collected with different genotyping platforms

BENEFITS TO INDUSTRY

- Large N<<P datasets
- Potential to play a major role in Precision Medicine Initiative
- Biometric prediction
- Improved power of clinical trials
- Identify subsets of patients that benefit from treatments that "fail" in clinical trials

Prediction of Complex Genetic Traits

Kirk Wilhelmsen

School of Medicine - Genetics

ISUCRC

INDUSTRY & UNIVERSITY COOPERATIVE RESEARCH PROGRAM

Genomic Estimated Breeding Values (GEBV) is big business

- Animal breeding schedules for increasing fat %, meat content , etc.
- Plant breeding to improve growth size, rates, water requirements, etc.
- Currently NOT regularly used in Human studies

- **Modeling Complexities**
  - N<<p
  - Genetic Architecture
  - Dominance
  - Relatedness
  - Epistasis
  - Gene-environment

e.g. Genetic Architecture

- **Polygenic**
  - Many common genes contribute a little to the overall GEBV
  - referred to as Linear Mixed Models
- **Large effect + polygenic**
  - Allows some genes to be important
  - Solutions based on Bayesian Sampling

Prediction of Complex Genetic Traits

Kirk Wilhelmsen

School of Medicine - Genetics

Cross Validation (CV) Prediction: Alzheimer's Dx

Example With Real Data

- Predictions of AD disease status using the BSLMM and 1508 cases and controls, 1,000,000 genetic markers
- % Training x % Test
- 50 replicates
- Includes Sex, APOE, Age, 10-PC (ethnicity) as Covariates

Future Plans

- We have capability to simulate massive realistic datasets
  - We sample from simulated DNA sequence 300,000 chromosome in 20mb segments
  - Calculate realistic phenotypes based on simulated genomes and exposure
- Compare and contrast prediction methods where the truth is known and sample size is not an issue