Spring 2018

Overview of ongoing & future research

Laura Balzer

Ongoing & Future Research

Laura B. Balzer, PhD

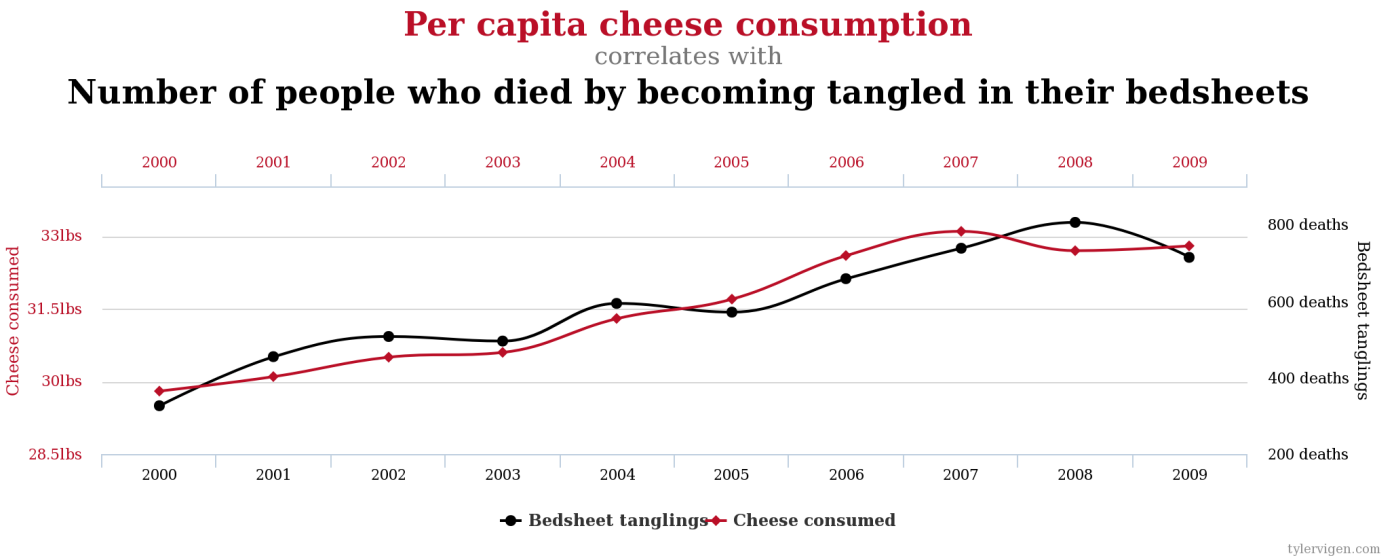
Department of Biostatistics & Epidemiology
School of Public Health & Health Sciences
University of Massachusetts, Amherst

Prospective Student Day
Feb 2018

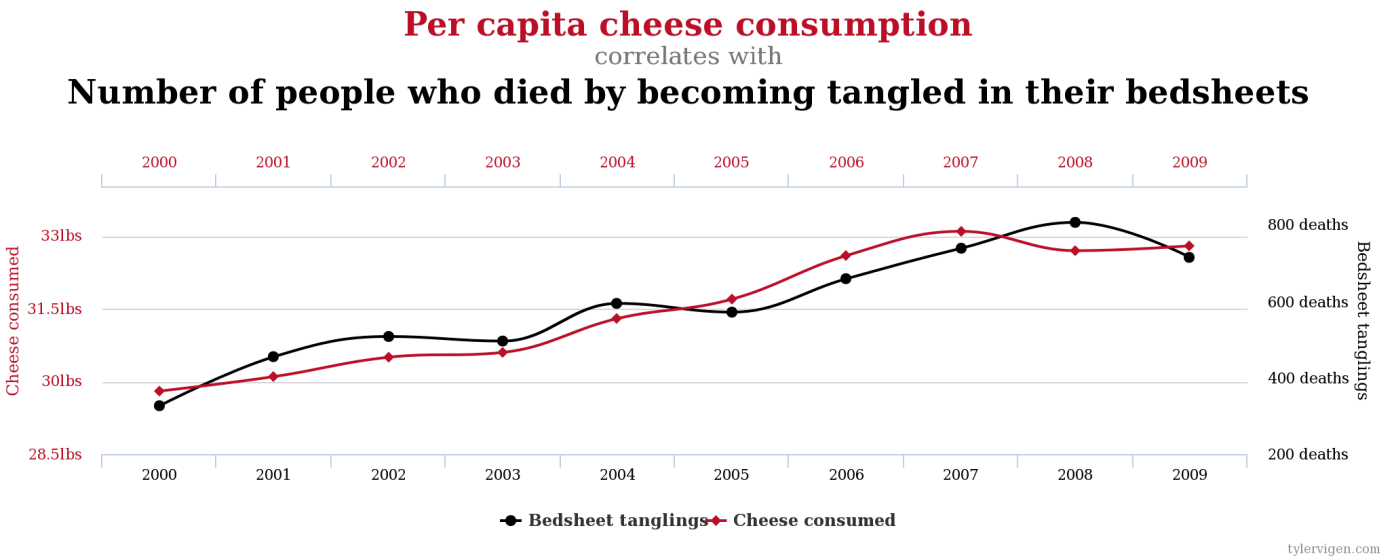
UMASS AMHERST

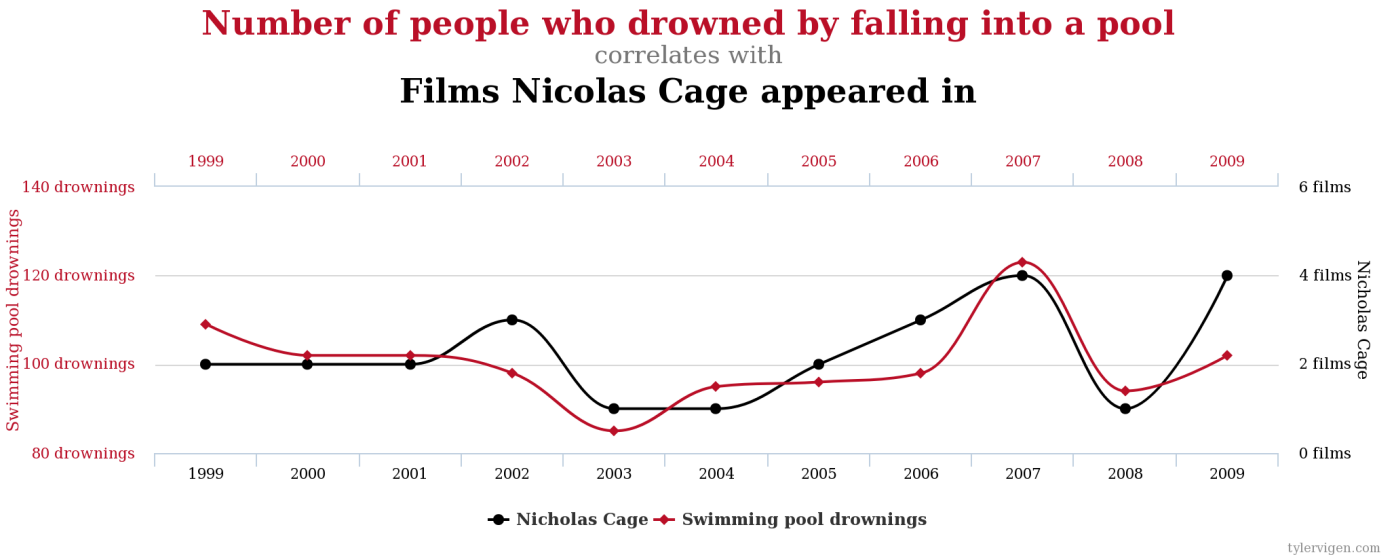
SEARCH
SUSTAINABLE EAST AFRICA RESEARCH IN COMMUNITY HEALTH

Causal inference



Causal inference





Causal inference: "Correlation is not causation"

Potential sources of association

(a) direct effects

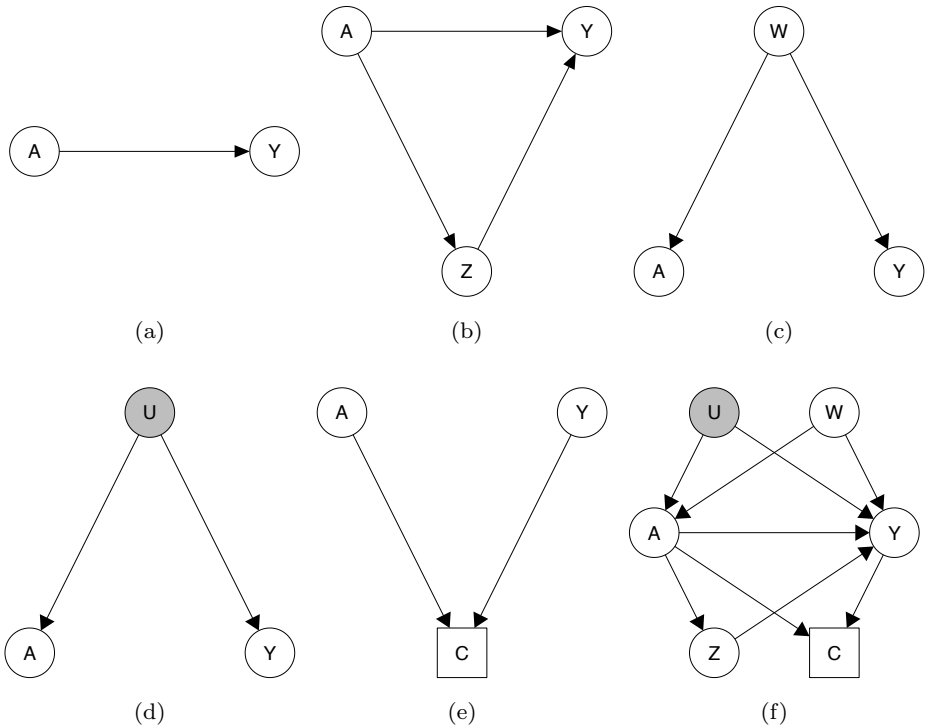
(b) indirect effects

(c) measured confounding

(d) unmeasured confounding

(e) selection bias

(f) all



Causal inference: a roadmap

0. Scientific question
1. Causal model representing **real** knowledge
2. Counterfactuals & causal parameter
3. Observed data & link to causal model
4. Identify: Knowledge + data sufficient?
5. Commit to an estimand as close to question as possible, and a statistical model representing real knowledge
6. Estimation
7. Interpretation

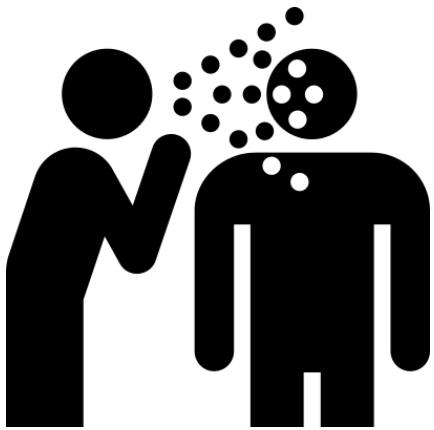
Want more? Biostat690B

Machine learning for prediction

- How can we build an optimal predictor for health outcomes?
  - "Optimal" depends on the research question
  - Single variable?
  - Parametric regression?
  - Too many variables; too little knowledge
- Machine learning: Automated approach to flexibly discover complex relationships from data
  - Super Learner: ensemble method to best weighted combination of algorithms

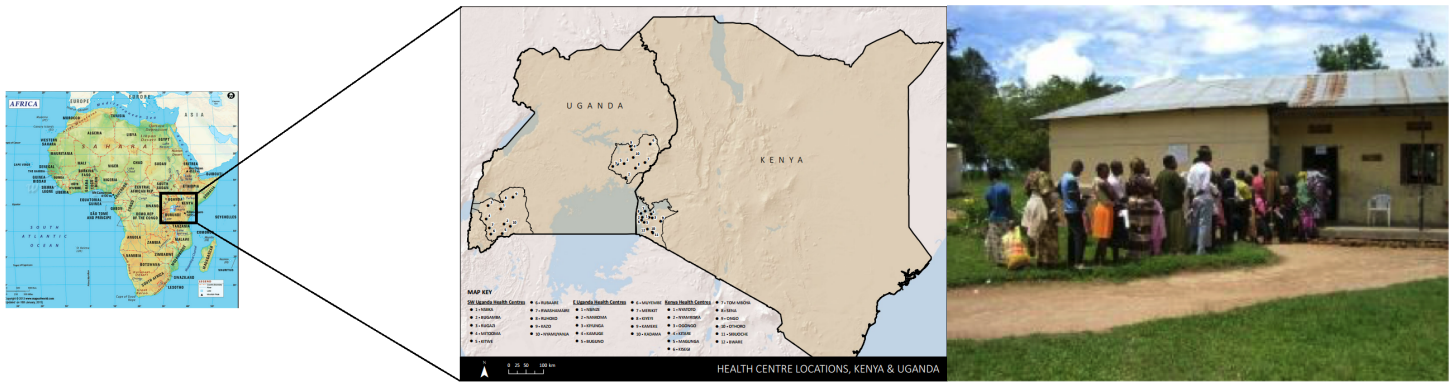
Statistics with highly dependent data

- Our health and well-being rarely are i.i.d. (independent identically distributed)
  - Clustering within households, neighborhoods, schools, communities
  - Longitudinal (repeated measures)
  - Hierarchical or multi-level
- Account for highly dependent data while minimizing assumptions
  - Let's not assume to know more than we actually know
    - Machine learning with statistical inference



The SEARCH Study - Integrating research & practice

- Ongoing cluster randomized trial (NCT:01864603)
  - Baseline: 2013-2014
- 32 communities in rural Uganda and Kenya
- >320,000 persons
- Pragmatic trial: the effect community-based HIV testing and ART-based interventions in real world settings



www.searchendaids.com

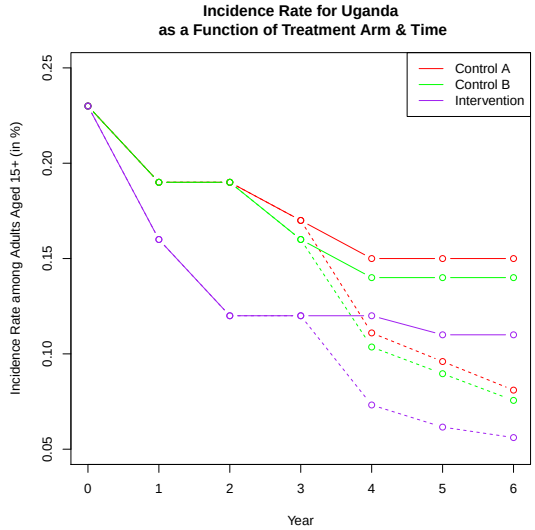
The SEARCH Study - Integrating research & practice

- Intervention:
  - Annual community-wide HIV testing
  - Patient-centered ART for all HIV+
- Control:
  - Baseline community-wide HIV testing
  - Country-guided ART for HIV+
- Primary outcome:
  - Three-year cumulative incidence of HIV
    - Proportion of baseline HIV-negative adult residents who become HIV-positive within 3 years
  - Cohort: 118,038 individuals

The SEARCH Study - Integrating research & practice

Design and analysis issues

- What is our parameter of interest?
  - How do we translate our scientific question into a causal quantity?
- How should we design the trial?
  - To pair-match or not to pair-match?
- How should we analyze our data?
  - Maximally unbiased and efficient estimator?
  - Statistical inference?
  - Adaptive yet pre-specified analysis?
- Ever-changing landscape of HIV prevention and treatment



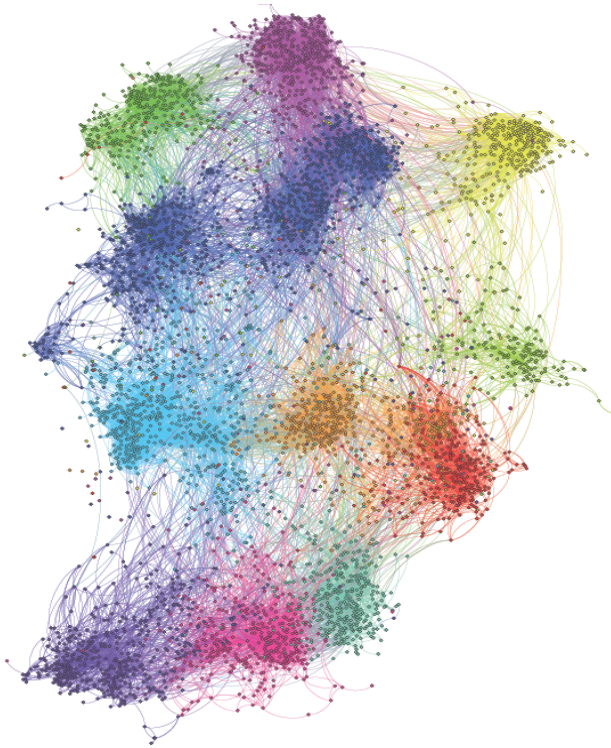
Machine learning in the field

Prediction: Who is at high risk for HIV infection?

- Self-assessment of risk
  - After community sensitization, education, counseling
- HIV-negative individual with a HIV-positive partner
- Who else?
  - Self-assessment may not be accurate
  - Basic demographic risk groups (e.g. young women) may miss many in a generalized epidemic setting
- Developed and deployed Super Learner on tablets
  - Risk classification for intervention targeting in real time and with limited resources

Social networks and public health

- Our health and well-being is affected by members of our social (and sexual) network
- How do we incorporate these dependencies into a causal and statistical inference framework?
- How can we use measures of the network to optimize interventions?
- How can we incorporate our knowledge of the network in our design and analysis?



***Community-wide social network**: Nodes (persons) are color-coded by village represent social links between persons.*

Thank you & Acknowledgements!


Supported by NIAID of the NIH , PEPFAR, Bill and Melinda Gates Foundation, and Gilead Sciences. The SEARCH project gratefully acknowledges the Ministries of Health of Uganda and Kenya, our research team, collaborators and advisory boards, and especially all communities and participants involved.