A data science framework for learning health systems

Yates Coley, PhD
Assistant Investigator
Kaiser Permanente Washington Health Research Institute
rebecca.y.coley@kp.org

October 23, 2018
Data!
Data!
How can we learn from all available data in order to improve health care?
Data!

How can we learn from all available data in order to improve health care?

… and how do we actually influence care delivery?
Learning Health Systems

Precision Medicine

Patient-Centered Care
What do statisticians need to know?
What do statisticians need to know?  

What role do statisticians play?  

Patient-Centered Care  

Precision Medicine
Learning Health Systems

Patient-Centered Care

Precision Medicine
Learning Health Systems

Precision Medicine

Patient-Centered Care

• Targeting treatment
Learning Health Systems

Precision Medicine

Patient-Centered Care

- Targeting treatment
- Risk prediction
Learning Health Systems

Precision Medicine

Patient-Centered Care

- Targeting treatment
- Risk prediction
- Population inference
Learning Health Systems

Precision Medicine

Patient-Centered Care

Preferences and priorities
Learning Health Systems

Precision Medicine

Patient-Centered Care
Learning Health Systems

Data systems to enable decision support
Statistical models to learn from health data
Dissemination via clinical decision support
Curation for improvement and adaptation
Non-statistical considerations
Data Systems to Enable Learning

• Electronic health record (EHR)
Data Systems to Enable Learning

• Electronic health record (EHR)
  • Captured as part of care process, not research
  • Contain data on encounters, diagnoses, procedures, prescriptions, vitals, lab results, provider notes
  • Collaborate with content-area experts, EHR users
  • Informative missing data—cross-section, longitudinal
  • Research-quality requires investment, commitment
  • Potential for wonderfully big datasets
Data Systems to Enable Learning

• Electronic health record (EHR)

• Supplemental data sources
Data Systems to Enable Learning

- Electronic health record (EHR)

- Supplemental data sources
  - Registries- immunization, disease, mortality
  - Geographic, environmental data
  - Supplemental research data collection
  - Prior research studies
Data Systems to Enable Learning

- Electronic health record (EHR)

- Supplemental data sources

- Clinical use, iterative process of LHS
Data Systems to Enable Learning

• Electronic health record (EHR)

• Supplemental data sources

• Clinical use, iterative process of LHS
  • Sustainable data source for continued monitoring
  • Retrospective analysis vs. real-time availability
Data systems to enable decision support

Statistical models to learn from health data

Dissemination via clinical decision support

Curation for improvement and adaptation

Non-statistical considerations
Statistical models to learn from health data
Statistical models to learn from health data

- Reflect sources of variability
Statistical models to learn from health data

• Reflect sources of variability
• Reflect traditional sources of bias
Statistical models to learn from health data

• Reflect sources of variability
• Reflect traditional sources of bias
• Reflect EHR-specific sources of bias
Statistical models to learn from health data

- Reflect sources of variability
- Reflect traditional sources of bias
- Reflect EHR-specific sources of bias
- Accommodate different sources of data
Data systems to enable decision support
Statistical models to learn from health data
Dissemination via clinical decision support
Curation for improvement and adaptation
Non-statistical considerations
Clinical decision support tool
Clinical decision support tool

• Integration with database and prediction model- no manual entry of patient data
Clinical decision support tool

• Integration with database and prediction model- no manual entry of patient data
• Just visualizing patient data powerful
Clinical decision support tool

• Integration with database and prediction model - no manual entry of patient data
• Just visualizing patient data powerful
• Graphic and text explanations
Clinical decision support tool

- Integration with database and prediction model - no manual entry of patient data
- Just visualizing patient data powerful
- Graphic and text explanations
- Electronic and/or print access to tool outside of clinic
Clinical decision support tool

- Integration with database and prediction model - no manual entry of patient data
- Just visualizing patient data powerful
- Graphic and text explanations
- Electronic and/or print access to tool outside of clinic
- Focus on outcomes of interest to patients and relevance to care plan
PSA/Biopsy Summary

- PSA
- Typical PSA Range for Men in Active Surveillance
- No cancer found
- Grade group 1 cancer found
- MRI

Graph showing age against PSA levels and MRI/Biopsy results.
If 100 men with a similar age, diagnosis, and PSA and biopsy history had their prostate surgically removed today, what cancer grade would be found?

- Grade group 1: 79
- Grade group 2: 18
- Grade group 3: 2
- Grade group 4-5: 2

Patient Last Name, FirstName
MRN: JH1234567
DOB: 01/01/00
Date: 05/08/2017
If 100 men like you had their prostates surgically removed today, after 5 years...

- **Would be cured**: 97
- **Would have PSA recurrence**: 3
- **Would have metastatic disease**: <1

*Note: Because of rounding to the nearest whole number, counts may not match grade group totals.*
Data systems to enable decision support
Statistical models to learn from health data
Dissemination via clinical decision support
Curation for improvement and adaptation
Non-statistical considerations
Curation for improvement and adaptation
Curation for improvement and adaptation

• Monitor impact
  • Document any intervention in EHR
  • Evaluate accuracy of estimates, predictions
  • Prevent harmful feedback loops
Curation for improvement and adaptation

• Monitor impact
  • Document any intervention in EHR
  • Evaluate accuracy of estimates, predictions
  • Prevent harmful feedback loops

• Accommodate impact of interventions
Curation for improvement and adaptation

- Monitor impact
  - Document any intervention in EHR
  - Evaluate accuracy of estimates, predictions
  - Prevent harmful feedback loops
- Accommodate impact of interventions
- Accommodate differences in population
Curation for improvement and adaptation

• Monitor impact
  • Document any intervention in EHR
  • Evaluate accuracy of estimates, predictions
  • Prevent harmful feedback loops

• Accommodate impact of interventions
• Accommodate differences in population
• Incorporating new data sources
Data systems to enable decision support
Statistical models to learn from health data
Dissemination via clinical decision support
Curation for improvement and adaptation

Non-statistical considerations
Non-Statistical Considerations

• Investment, enthusiasm from clinicians
• Patient and user stakeholders
• Interdisciplinary team
• Project management
• Time and funding
• Incentive structure
• Institutional support
• IRB strategy
Acknowledgments

This research was supported by the *Patrick C. Walsh Prostate Cancer Research Fund* and a Patient-Centered Outcomes Research Institute (PCORI) Award (ME-1408-20318).

“The statements presented in this presentation are solely the responsibility of the authors and do not necessarily reflect the views of the Patient-Centered Outcomes Research Institute (PCORI), its Board of Governors or Methodology Committee.”

Additional support by the *NIMH* (U19 MH099201) and the Learning Health Systems program at *Kaiser Permanente Washington*

Presentation content also contributed by biostatistics investigators at Kaiser Permanente Washington Health Research Institute: *Jen Nelson, Jennifer Bobb, Andrea Cook*, and *Susan Shortreed*