

Teaching Bioinformatics Students about Clinical Data

Ted Laderas

2019-04-19

Introduction: Ted Laderas

- Assistant Professor
 - Division of Bioinformatics and Computational Biomedicine
 - Department of Medical Informatics and Clinical Epidemiology
 - Oregon Health & Science University
- Research Interests
 - Education (active learning)
 - Interactive Visualization
 - Immunoinformatics
 - Drug Sensitivity/Resistance in Cancer

Overview

- Why teach bioinformatics students about clinical data?
 - Encourage translational research
- How do we teach them?
 - Clinical Data Wrangling
 - Data Analytics
 - BioData Club
- Who are examples of our success?
 - Gabby Choonoo
 - Connor Smith

Motivation: Why?

- Research is multi-disciplinary
- Need for Translational Research Workforce
- Bridge gaps between computational biology and clinical realms

Bioinformatics Students: Clinical Data Opportunities

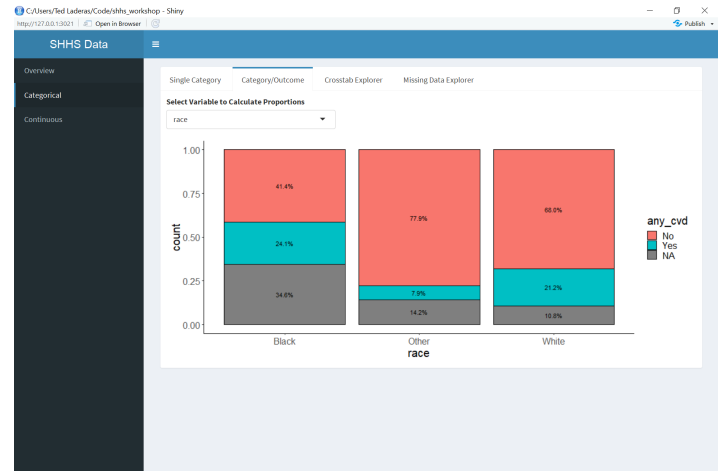
- Clinical Data Wrangling Workshop
- BMI569 Data Analytics
- BioData Club

Clinical Data Wrangling

- Designed for both clinical and bioinformatics students
- Introduce students to data quality issues with clinical data
- Use Sleep Heart Health Study data
 - Partnership with <http://sleepdata.org>
- http://laderast.github.io/clinical_data_wrangling/

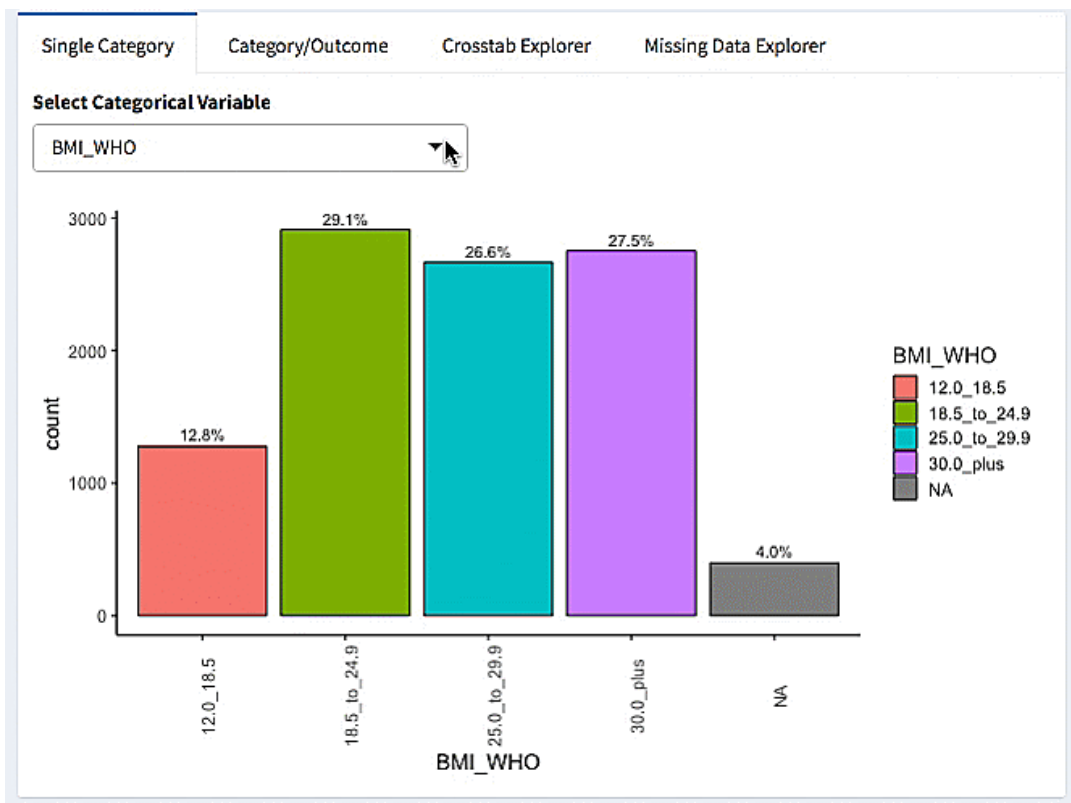
The Goal: build a predictive model of Cardiovascular Disease

- Biology of Sleep
- Data Quality Issues
- Exploring the Dataset
- Building a predictive model
- Reporting your predictive model



Making Decisions by Exploring Data

- Shiny app (burro) that lets students explore dataset
- Code-less, assumes no coding expertise



Building Predictive Models by Making Decisions

Step 1: Initial Model

Choose your initial model from the following. You don't need to show any figures here.

- `any_cvd` (your outcome)
- `age_s1`
- `gender`
- `bmi_s1`
- `neck20`

Hide

```
#show your code for the basic model here
```

Step 2: Do you add `race` to your model?

Put a short definition of `race` from your model. If you think it is important to add `race` and you are satisfied with the quality in the dataset, show a figure here. If you don't think it's important or you aren't satisfied with the quality of the `race` variable, show a figure here.

Hide

```
#put model code here
```

Outcomes

- Currently assessing impact on students

BMI569: Data Analytics

- Summer Hybrid Course taught jointly with Kaiser Permanente
- Focus on helping make decisions in a Health Care Setting
- Sociotechnical challenges in implementing a predictive model
 - Implementing a predictive model (R-based)
 - Organizational Lectures and Challenges from Kaiser

David Dorr, Tracy Edinger, *Ted Laderas*, Shannon McWeeney, *Delilah Moore*, and *Brian Sikora*

BMI569: Practical Coursework

- 8 weeks online coursework in R/SQL and 1 week in person
- Patients assess and query a simulated data warehouse
- Rmarkdown document based assignments
 - provide guidance
- <http://laderast.github.io/AnalyticsCourse>

Mixing Clinical and Bioinformatics

- Roughly 50% Clinical/50% Bioinformatics
- Opportunity! Pair them up for joint learning
- Learn skills and challenges from each other

Goals

- Predict 30-day hospital readmissions
- Joining/Querying/Data
- Implementing the LACE metric and assessing it
- Kaiser Experience in Implementation and Impact
- Communicating your results to an executive team

Length of stay (days)	Score (circle as appropriate)
1	1
2	2
3	3
4-6	4
7-13	5
14 or more	7

Step 2. **Acuity of Admission**
Was the patient admitted to hospital via the emergency department?
If yes, enter "3" in Box A, otherwise enter "0" in Box A

Step 3. **Comorbidities**

Condition (definitions and notes on reverse)	Score (circle as appropriate)
Previous myocardial infarction	+1
Cerebrovascular disease	+1
Peripheral vascular disease	+1
Diabetes without complications	+1
Congestive heart failure	+2
Diabetes with end organ damage	+2
Chronic pulmonary disease	+2
Mild liver or renal disease	+2
Any tumor (including lymphoma or leukemia)	+2
Dementia	+3
Connective tissue disease	+3
AIDS	+4
Moderate or severe liver or renal disease	+4
Metastatic solid tumor	+6
TOTAL	

If the TOTAL score is between 0 and 3 enter the score into Box C.
If the score is 4 or higher, enter 5 into Box C

Step 4. **Emergency department visits**
How many times has the patient visited an emergency department in the six months prior to admission (not including the emergency department visit immediately preceding the current admission)? _____
Enter this number or 4 (whichever is smaller) in Box E

<https://greatplainsqin.org/wp-content/uploads/2015/01/Lace-Index-Scoring-Tool.pdf>

Analytics Course: Successes

- Nominated for a course award by our students
- Students have gone on to work for Kaiser

BioData Club

- Lifelong Learning Group about Data Science
- Cross Disciplinary / Hierarchy Flattening
- Psychologically safe space to learn
- Teaching Opportunity
- Lots of workshops/social activities!



<http://biodataclub.github.io>

BioData Club Kit

- Start a BioData Club at your Institution!
- Part of CD2H (Center for Data to Health)
- Contact me if interested!



NATIONAL CENTER
FOR DATA TO HEALTH

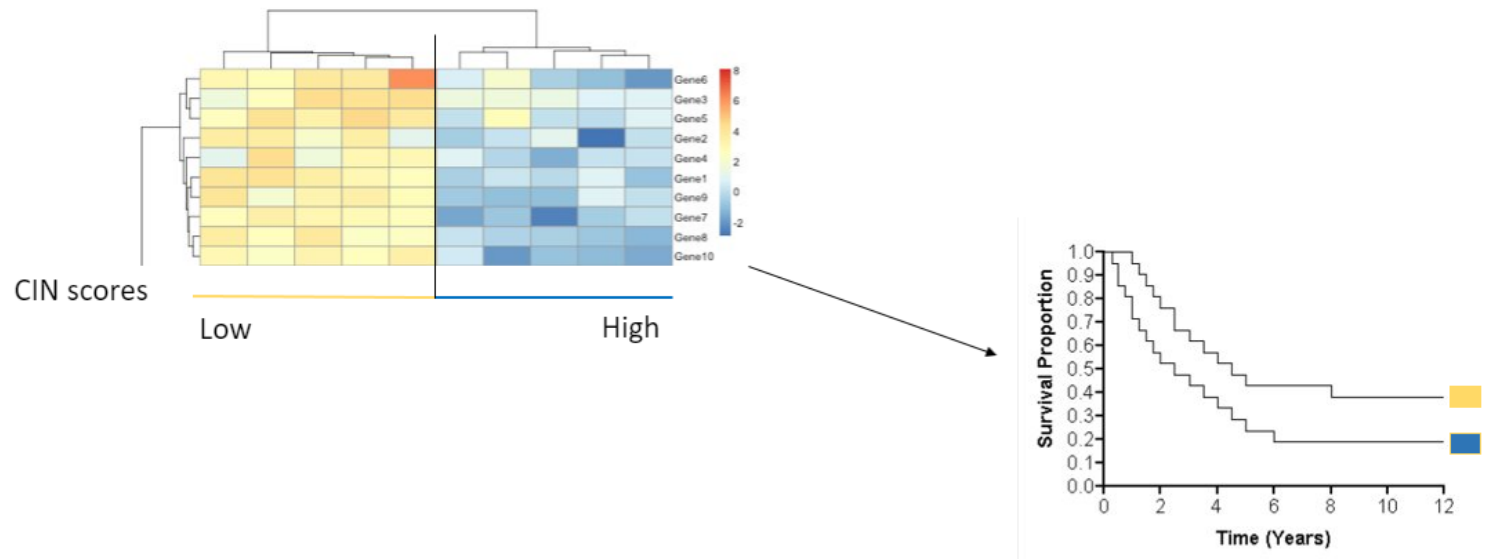
Student Successes

- Gabrielle Choonoo
- Connor Smith

Gabrielle Choonoo

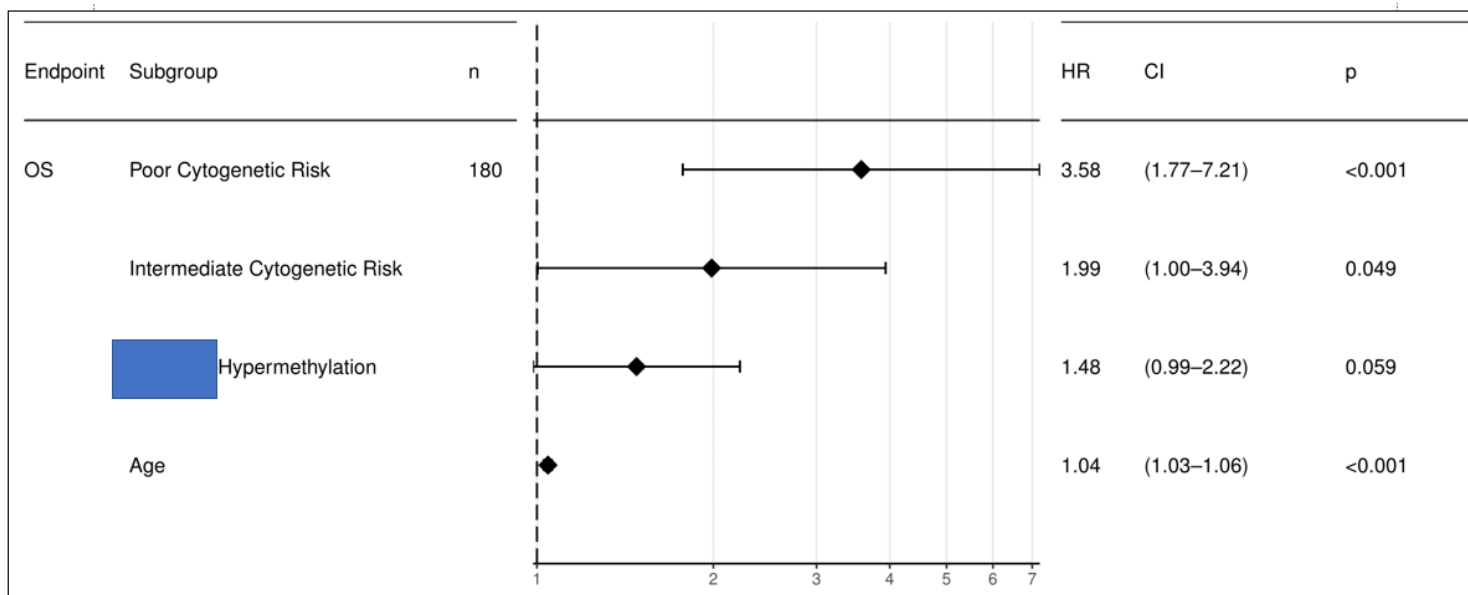
"Genomic Stability of Solid Tumors and Clinical Applications"

Investigate the predictive ability of the genomic instability annotations across a variety of cancer types



Connor Smith

- OSU Graduate in Biology
- Thesis: Methylation Effects in AML
- Now working in a clinical setting!



Principles and Values

- Reduce barriers between clinical and computational cultures
- Pair clinical/bioinformatics students for co-learning
- Flatten hierarchies for co-learning
- Value lifelong learning as an attitude

Funding Acknowledgements:

- T15 NLM Training Supplement Grant: 5T15LM007088
- NCATS Center for Data to Health (CD2H) Grant: U24TR002306

Keep in Touch!

laderast@ohsu.edu | @tladeras | <http://laderast.github.io>

- <http://laderast.github.io/AnalyticsCourse>
- <http://laderast.github.io/burro>
- http://laderast.github.io/clinical_data_wrangling/
- BioData-Club Kit: <http://biodataclubkit.github.io>

Learn more about our graduate program!

- 4/23, 10 AM - 2 PM @ OSU PreHealth Fair (Memorial Union Ballroom)
- Webinars about
 - Clinical and Bioinformatics Majors
 - email Lauren Ludwig: ludwigl@ohsu.edu