

#### **VCU Biostatistics Information Session**

Leroy Thacker, Ph.D. Professor Department of Biostatistics



School of Population Health

#### What is Biostatistics?

• Statistics applied to the collection, analysis, and interpretation of biological data and especially data relating to human biology, health, and medicine – Meriam-Webster

 Biostatistics is the discipline concerned with how we ought to make decisions when analyzing biomedical data. It is the evolving discipline concerned with formulating explicit rules to compensate both for the fallibility of human intuition in general and for biases in study design in particular. – Berger, & Rosser Matthews, 2006



## **Biostatistics in Public Health**

- Biostatistics touches every facet of Public Health
  - Collaboration
    - Cancer
    - Chronic Diseases
    - Environment
    - Statistical Genetics
    - Human Growth & Development
  - Methodological development
    - New statistical methods/study designs



## VCU Department of Biostatistics

- We are housed in the VCU School of Population Health (SOPH)
  - The SOPH was formed in July, 2023 with the expectation of becoming an approved School of Public Health
- VCU Biostatistics has been around a bit longer than the SOPH!
  - 1958 : Division of Biometry forms in the Medical College of Virginia (MCV)
  - 1969: VCU formed from MCV RPI merger; began recruitment of MS and PhD students
  - 1972: Officially changed names to the Department of Biostatistics
  - 2013: Introduced the PhD and MS in Genomic Biostatistics
  - 2023: Joined the VCU School of Population Health as a founding department



#### VCU Department of Biostatistics







#### VCU Biostatistics Faculty Members

- There are 17 full-time faculty members in the department
  - The areas of expertise and methodological development for these faculty covers all areas of biostatistics in public health
- The next few slides highlight some of our faculty members and their areas of expertise



#### Sensor & Metabolic Chamber Data

- Shanshan Chen, PhD
- Ekaterina Smirnova, PhD



School of Population Health





#### A Generic Algorithm for Sleep-Wake Cycle Detection using Unlabeled Actigraphy Data

Shanshan Chen, Robert Perera

Department of Biostatistics Virginia Commonwealth University Richmond VA, USA {shanshan.chen, robert.perera}@vcuhealth.org

Matthew M. Engelhard, Jessica R. Lunsford-Avery, Scott H. Kollins Department of Psychiatry Duke University

Durham NC, USA {m.engelhard, jessica.r.avery, scott.kollins}@duke.edu

Bernard F. Fuemmeler

Department of Health, Behavior and Policy Virginia Commonwealth University Richmond VA, USA Bernard.Fuemmeler@vcuhealth.org Data from wearables can predict risk of death in older adults



Researchers analyzed health data from wearable technology trackers to predict the five-year mortality risks of the U.S. aging population. (Getty Images)

By Alex Nowak University Public Affairs (804) 828-7028 alnowak@vcu.edu

Wednesday, Oct. 30, 2019

In 2019, wearable technology is ubiguitous.

Whether used for tracking sleep patterns, monitoring heart rate or logging exercise, nearly 57 million U.S. adults - roughly a quarter of the adult population will use a wearable device such as a fitness tracker or smartwatch at least once a month this year, according to an eMarketer study. Moreover, wearables are becoming popular among Americans 55 and older. Their use is projected to increase more than 15 percent since 2018.

Noting the upswing in wearable technology use and Americans' declining activity levels as they age, Ekaterina Smirnova, Ph.D. assistant professor in the Department of Biostatistics in the VCU School of Medicine, and co-leaders of Johns Hopkins University's Wearable and Implantable Technology group - Ciprian Crainiceanu, Ph.D., and Vadim Zipunnikov, Ph.D. - studied how physical activity levels can predict death.

Ekaterina Smirnova, Ph.D.



#### **Computational Genomics**

- Mikhail Dozmorov, PhD
- Jinze Liu, PhD
- Katarzyna Tyc, PhD





#### Spatial Data Analysis

- David Wheeler, PhD
- Dipankar Bandyopadhyay, PhD



SMARTp: A SMART design for nonsurgical treatments of chronic periodontitis with spatially referenced and nonrandomly missing skewed outcomes

Jing Xu<sup>1,2</sup> Dipankar Bandyopadhyay<sup>3</sup> Sedigheh Mirzaei Salehabadi<sup>4</sup> Bryan Michalowicz<sup>5</sup> Bibhas Chakraborty<sup>1,6,7</sup>



FIGURE 1 CAL data Note: Panel (A) shows the observed CAL for a patient with a missing incisor, where the shaded boxes represent teeth, the circles represent sites, and gray lines represent neighbor pairs that concat adjacent sites on the same tooth and sites that share a gap between teeth. "Gap" in the figure indicates, for example, the four sites in the gap between teeth #4 and 5. The tooth numbers are indicated, and excludes the four third molans: 1, 16, 17, 32. The vertical and horizontal lines separate the mouth into four quadrants, with the molars (#2 and 3, #14 and 15, #18 and 19, #03 and 31), premolars (#4 and 5, #12 and 13, #20 and 21, #28 and 29), canine: (#6, 12, 22, 27) and incisors (#7-10, #23-26). Panel (B) presents the frequency density plot of the CAL (conded to the nexest millimeters) for each tooth type from the GAAD dataset

Spatial Analysis of the Impact of a School-Level Youth Violence Prevention Program on Violent Crime Incidents in the Community

Saba W. Masho<sup>1,2</sup> • Keith W. Zirkle<sup>3</sup> • David C. Wheeler<sup>3</sup> • Terri Sullivan<sup>2,4</sup> • Albert D. Farrell<sup>2,4</sup>



Fig. 1 Estimated relative risk of youth violence for 10- to 18-year-old offenders by census block group by year. The block groups outlined in green and purple outline middle school zones 1 and 2, respectively. Values below 1 (shades of blue) indicate a decreased risk



#### **Clinical Trial Design**

Randomization

a)

Adaptive Design

- Nolan Wages, PhD
- Roy Sabo, PhD
- Robert Perera, PhD





1.5-fold 2-fold 3-fold

100 125 150 175

b)

75 100 Patient

#### Image Analysis

• Nitai Mukhopadhyay, PhD











## Missing Data

• Yongyun Shin, PhD



#### Longitudinal Latent Variable Models Given Incompletely Observed Biomarkers and Covariates

Chunfeng Ren and Yongyun Shin Department of Biostatistics, Virginia Commonwealth University, Richmond, Virginia 23219

$$\begin{bmatrix} Y_{1i} \\ Y_{2i} \end{bmatrix} = \begin{bmatrix} X_{1i} & 0 \\ 0 & X_{2i} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} Z_{1i} & 0 \\ 0 & I_{p_2} \end{bmatrix} \begin{bmatrix} b_{1i} \\ b_{2i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1i} \\ 0 \end{bmatrix} + \begin{bmatrix} a_{1i} + c_{1i} \\ 0 \end{bmatrix},$$
(7)

$$\begin{aligned} & for Y_{1i} = \begin{bmatrix} R_i \\ S_i \end{bmatrix} X_{1i} = \begin{bmatrix} I_{J \times k_i} & W_{ui} \otimes I_J & 0 \\ 0 & 0 & W_{si} \end{bmatrix} \beta_1 = \begin{bmatrix} 1_{k_i} \otimes \gamma_0 \\ \beta_u \otimes \gamma_1 \\ \beta_s \end{bmatrix} \\ & Z_{1i} = \begin{bmatrix} B_i \otimes I_J & 0 \\ 0 & 1_{k_i} \otimes I_{p_1} \end{bmatrix} b_{1i} = \begin{bmatrix} b_{ui} \otimes \gamma_1 \\ b_{si} \end{bmatrix} \varepsilon_{1i} = \begin{bmatrix} \varepsilon_{ui} \otimes \gamma_1 \\ \varepsilon_{si} \end{bmatrix} a_{1i} = \begin{bmatrix} 1_{k_i} \otimes a_i \\ 0 \end{bmatrix} a_{1i} d_{1i} = \begin{bmatrix} e_{i} \\ 0 \end{bmatrix} a_{1i} e^{\operatorname{var}(b_{1i}, b_{2i})} = \begin{bmatrix} \tau_{11} & \tau_{12} \\ \tau_{12}^T & \tau_{22} \end{bmatrix}, \end{aligned}$$



## VCU Biostatistics Programs of Study

- We offer two programs of study and a total of five tracks within the two programs
  - PhD
    - Traditional Track
    - Optional Genomics Concentration
  - MS
    - Traditional Track
    - Genomics Concentration Track
    - Clinical Research Concentration



#### VCU Biostatistics Coursework

- All Programs
  - Biostatistical Computing
  - Mathematical Statistics I
  - Analysis of Biomedical Data I and II
  - Clinical Trials
  - Consulting
  - Seminar
  - Scientific Integrity



### VCU Biostatistics Coursework

- The traditional and genomics MS track also require:
  - 12 hours of elective credits
  - Genomics track also requires Fundamentals of Molecular Genetics (or other relevant course) Statistical Methods for High-Throughout Genomic Data I and II (which reduces the number of elective hours)
- The PhD Program further requires:
  - Mathematical Statistics II
  - Biostatistical Methods I and II
  - Advanced Inference
  - Mixed Models and Longitudinal Analysis\*
  - Multivariate Analysis^
  - Survival Analysis
  - Statistical Methods for High-Throughout Genomic Data I and II<sup>^</sup>

\*Required only for traditional Biostatistics Track

^Required only for Genomics Track



## **Qualifying Examinations**

#### Applied Exam MS and PhD Program

**Courses:** Biostatistical Computing, Clinical Trials, Analysis of Biostatistical Data I and II

#### Timing:

End of summer following completion of program courses (typically end of first year)

#### Learning Goals Assessed:

- Application of Statistical Methods
- Computational Skills
- Written Communication Skills

#### Exam Type:

- Take-Home, Open-Book, 1-Week
- Set by Assessment Committee
- Focuses on single contextual "topic"
- IMRaD-Style write-up of methods and results
- Writing and coding taken into account

## **Qualifying Examinations**

#### Theoretical Exam

#### PhD Only

#### **Courses:**

Mathematical Statistics I and II, Biostatistical Methods I and II

#### Timing:

Summer Year 2

#### Learning Goals Assessed:

Theoretical and Methodological Acumen

#### Exam Type:

- In-Class and Timed
- Set by course instructors
- Similar to problems from class





## **Dissertation Proposal Defense**

- Timing:
  - Begin as soon as possible after Applied and Theoretical Qualifiers are passed
  - End no later than December of 4<sup>th</sup> year (3.5 academic years)
- Components:
  - Written proposal
    - Grant Proposal (F31)
    - Dissertation-Style Manuscript (no-page limit)
  - Oral defense to dissertation committee
    - Oral presentation (with slides) followed by Q&A from the committee



## **Dissertation Defense**

- Components:
  - Original Research
    - Methodological research in biostatistics, genomics, bioinformatics
    - Faculty mentorship
  - Written Dissertation
    - Usually consists of three or more publishable papers
    - At least one must be submitted for peer review
  - Oral Defense
    - Part 1: Public presentation: ~45 minute presentation/~15 minute discussion
    - Part 2: Private defense: Committee members ask questions

Timing: When dissertation research *is completed* – usually end of 5<sup>th</sup> year



# Why VCU?

- The Experiential Learning Emphasis
  - Didactics
  - Biostatistical Consulting Laboratory (BCL)
  - Student Summer Training Program







## Didactics

• The coursework balances standard didactic learning with the development of professional skillsets

PROVIDING CONSULTING EXPERIENCES THROUGH ROLE PLAYING IN A GRADUATE STATISTICS COURSE

R.T. SABO<sup>1</sup>

Virginia Commonwealth University

- Introduces other important aspects of being a professional biostatistician
  - Consulting
  - Manuscript review
  - Scientific writing and presentation
  - Seminars



#### **BIOS 603 Sequence**

- The biostatistical consulting sequence includes three different sections
  - Section 1 Academic Writing (Basics of writing techniques, scientific writing, and revision)
  - Section 2 Consulting Laboratory
  - Section 3 Manuscript Review (learn the peer review process by reviewing preprinted or medical statistical papers)



## **Biostatistical Consulting Laboratory**

- Hands on consulting experience
  - In-person meetings with clinical investigators
  - Small group presentations and team learning
  - Design, analysis and coding experience
  - Report generation
- BCL TAs
  - Supervises projects
  - Engages in administrative work and outreach





#### Summer Student Training Program

- Students are paired with faculty mentor to work on applied or methodological project
- Often used as precursor to dissertation research











# Funding

- Currently, we do not provide funding for MS students
- For PhD students, qualified applicants are invited to an open house early in the spring semester
  - Within a week, applicants will be notified if they have been offered admission with funding
  - Funding includes:
    - Tuition and fees (Currently \$46,563)
    - Health Insurance
    - Laptop
    - Annual Stipend (Currently \$34,000)



## After Graduation

- Some of the Employers of our recent graduates include:
- Proctor & Gamble
- Pfizer
- GSK
- Abbvie
- Siemens
- PharPoint
- Emmes

- FDA
- United Network for Organ Sharing (UNOS)
- US Military Academy
- Medical University of South Carolina (MUSC)
- University of Alabama-Birmingham (UAB)
- VCU Health

#### **Admission Requirements**

- For Both the MS and PhD Programs, Prerequisites Include:
  - Calculus and Multivariable Calculus
  - Linear Algebra
  - Concepts of Statistics
  - Introduction to Probability theory
- The GRE is Suggested for MS Applications and Required for PhD Applications



## How to Apply

- Go to the following link or use the QR code:
  - <u>https://admissions.vcu.edu/apply-to-vcu/</u>
- Feel Free to Browse our website:
  - <u>https://biostatistics.vcu.edu/</u>
- If you have questions, feel free to contact: Beth Ann Howard Education Coordinator (804) 827-2049 howardea@vcu.edu





#### Discussion



Visualization of Collatz Conjecture suggested by Edmund Harriss

