From Classroom to Clinical Impact: My Early Career as a Cancer Research Biostatistician

Lessons Learned and Daily Realities

Brad Loeffler, MS

The University of Iowa's Holden Comprehensive Cancer Center

April 11, 2025

Overview

- Educational & Professional Journey
  - Introduction and background
  - Why biostatistics?
  - Challenges and lessons
- Biostatistics at Holden Comprehensive Cancer Center
  - How I support the HCCC mission
  - Day-in-the-life look
  - Daily realities of academic cancer research
  - Tools and methodology used
  - Insights from my first five years

Educational & Professional Journey

Introduction

- Originally from Indianapolis, IN
- **2018** – BS in Actuarial Science from Indiana University-Purdue University-Indianapolis (IUPUI)
- **2020** – MS in Biostatistics from University of Iowa
- Biostatistician at the University of Iowa's Holden Comprehensive Cancer Center (HCCC) since **2020**

My Educational Journey

Actuarial Science major coming into undergrad

Failed exam FM 2/2 times, didn't enjoy actuarial science

Attended IUPUI's Biostatistics Discovery Day and applied for grad school!

Accepted to University of Iowa, moved to Iowa City

Worked as a statistics tutor. Interned under statistics professor.

First year of graduate school. Teaching assistant for introduction to biostatistics/epidemiology

Second year of graduate school. Preceptorship in machine learning for head and neck cancer patients. Research assistant at CPHS and HCCC.

Masters Statistician Intern at Eli Lilly and Company. Manufacturing department & cold-chain management

Hired remotely at the HCCC during COVID.

Why Biostatistics?

- **Biostatistics**: A branch of statistics concerned with analyzing outcomes from biological and health-related data.
- Mixes the traditional training from a statistics program with courses in public health, epidemiology, and clinical trial design
- Always enjoyed math, but my favorite parts of my actuarial science classes were the statistics components
- Several health events in my family influenced me to work in medicine, but I didn't want to be a doctor
- Considered public health, but it didn't have enough math!

Insights and Lessons Learned

- Teaching and tutoring were some of my most valuable resume-builders
- Eli Lilly and Company internship obtained in large part due to my interview and communication skills. My resume was not impressive!
- Networking is more about building healthy working relationships and proving your skills
  - Statistics tutoring -> Interning with statistics professor
  - Teaching assistant -> Center for Public Health Statistics
  - Preceptorship -> Research assistant -> Full-time employment at HCCC
  - Connected with alumni while working at Lilly

Insights and Lessons Learned

- School and internships are about adding as many tools to your statistical toolbox as you can
  - Eli Lilly: Bayesian methods, inverse transform sampling, kernel density estimation, Git, tidyverse packages, bootstrapping
  - Preceptorship: Machine learning methods, model selection, feature reduction
  - Center for Public Health Statistics: Mixed effects models, PROC SQL
  - HCCC: logistic regression, Cox regression
  - Consulting internship: Visual Basic, t-tests and chi-square tests
  - Teaching and tutoring
- Prove that you can do it in real life!

Interview Questions

- What software do you use the most? Why?
- What is an interesting coding challenge you've come across?
- How do you communicate with someone who isn't in your field?
- What kinds of models do you know about?
- What do you do when you get stuck on a statistical problem?
- Knowledge checks:
  - What is a p-value?
  - When do you use Fisher's exact test over a chi-square test?
  - What model would be most appropriate for XYZ outcome?
- Excited most about:
  - Experience teaching statistical concepts
  - Coding flexibility between R and SAS
  - Could speak to specific experience using specific methods

Biostatistics at Holden Comprehensive Cancer Center

The Holden Comprehensive Cancer Center

- The state of Iowa's only NCI-designated Comprehensive Cancer Center
- Investigators from over 200 medical specialties across 41 departments.
- Mission of the HCCC is to decrease the pain and suffering caused by cancer in Iowa, the surrounding community, and beyond.
- Interdependent missions of **research**, clinical care, and education to improve cancer prevention and treatment.

The Biostatistics Core

- Research at HCCC is conducted across a wide array of Shared Resources/core facilities

- Biospecimen Procurement and Molecular Epidemiology Resources (BioMER)
- **Biostatistics Core**
- Central Microscopy Research Facility
- Flow Cytometry Core
- Genomics Core
- Human Immunology Core [Dev]
- Microbiome Core [Dev]
- Population Research Core
- Radiation and Free Radical Research Core
- Viral Vector Core
- Imaging Core
- Proteomics Core

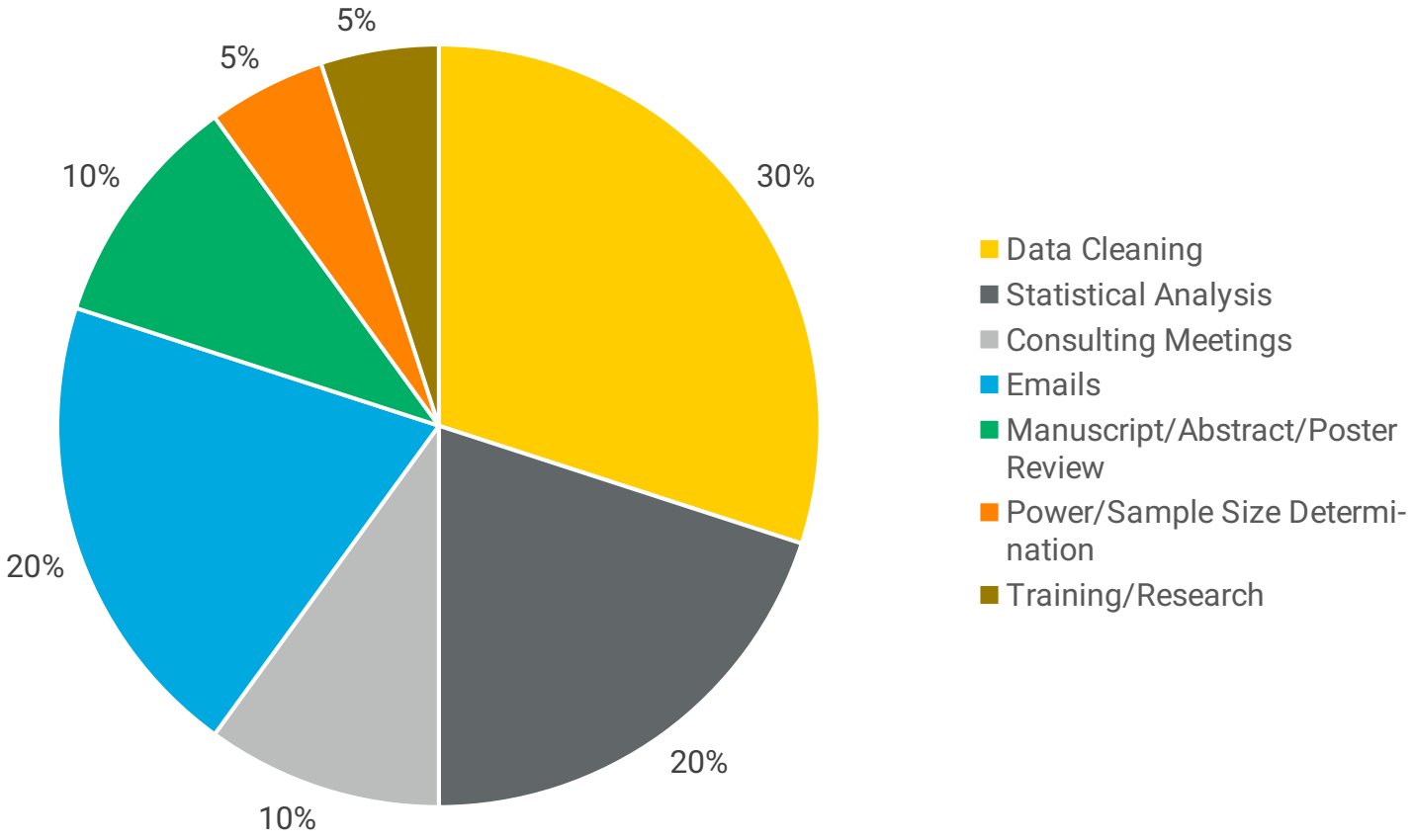
The Biostatistics Core

- The purpose of the Biostatistics Core is to provide statistical support for HCCC investigators in the design, analysis, and reporting of cancer research projects.
- Services include:
  - Consultation on study design, selection of outcome variables, and formulation of hypotheses
  - Specification of appropriate methods of data analysis
  - Sample size estimation
  - Protocol development
  - Generation of randomization schedules
  - Data analysis
  - Assistance in preparation of manuscripts
  - Education and training

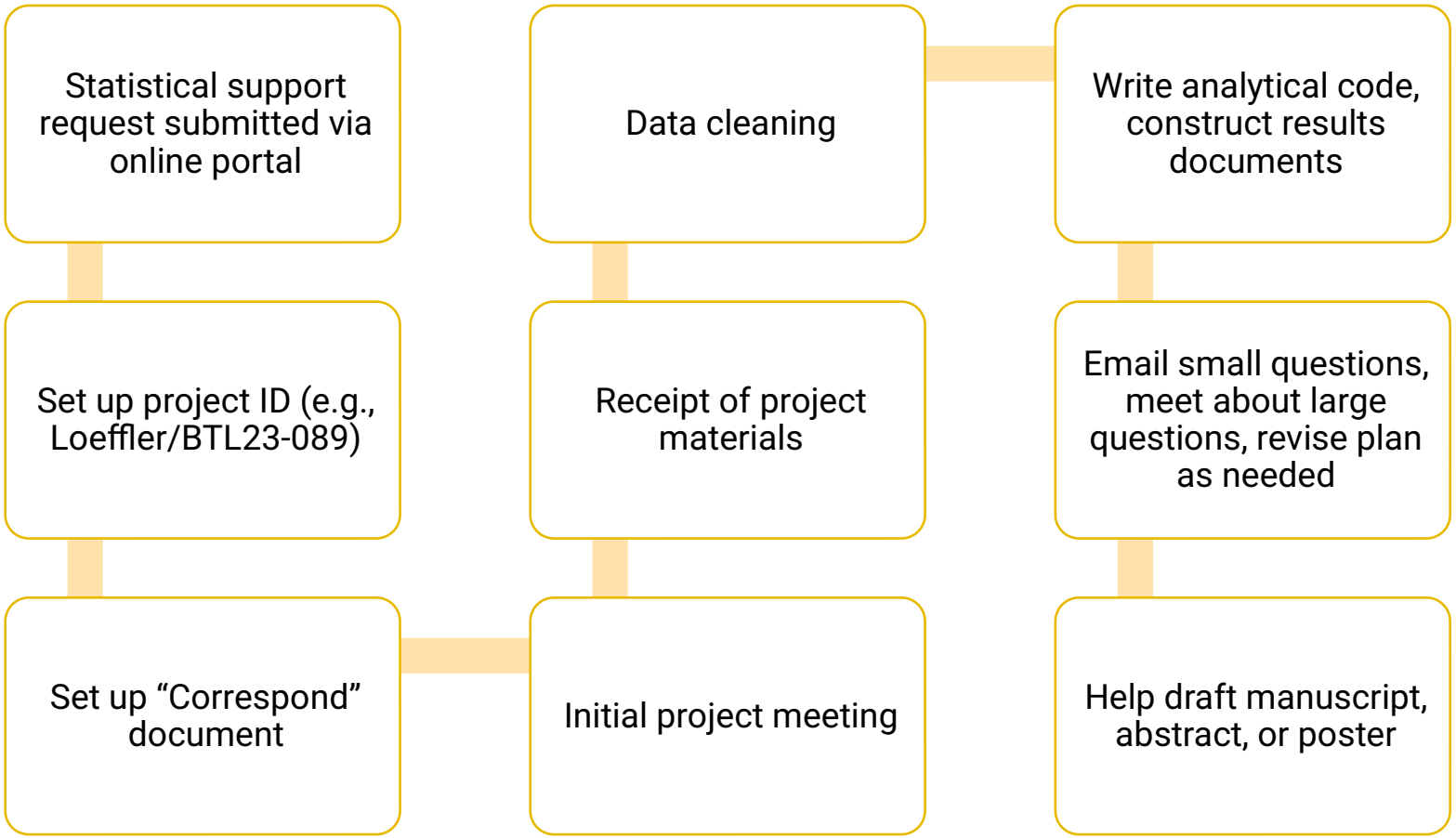
How I Spend My Time

- As a student, there was a disconnect between the between **big picture responsibilities** and **daily realities** of biostatistics
- What do you actually do day-to-day?
  - You get to work…
  - You put your lunch away…
  - You fill up your water bottle/coffee…
  - You check your email…
  - And then… what?

How I Spend My Time



Project Management (Retrospective Research)



Research Aims: From Idea to Reality

- Diverse set of investigators at various stages of project development
- **Example**: “I want to collect all patients at our institution with disease X who received treatment A, to see if treatment A helps patients.”
  - “Helps patients” compared to what?
  - How do you measure if a patient is “helped?”
  - What other treatments could a patient have gotten before A?
  - What patient or disease characteristics are also indicators of patient improvement?

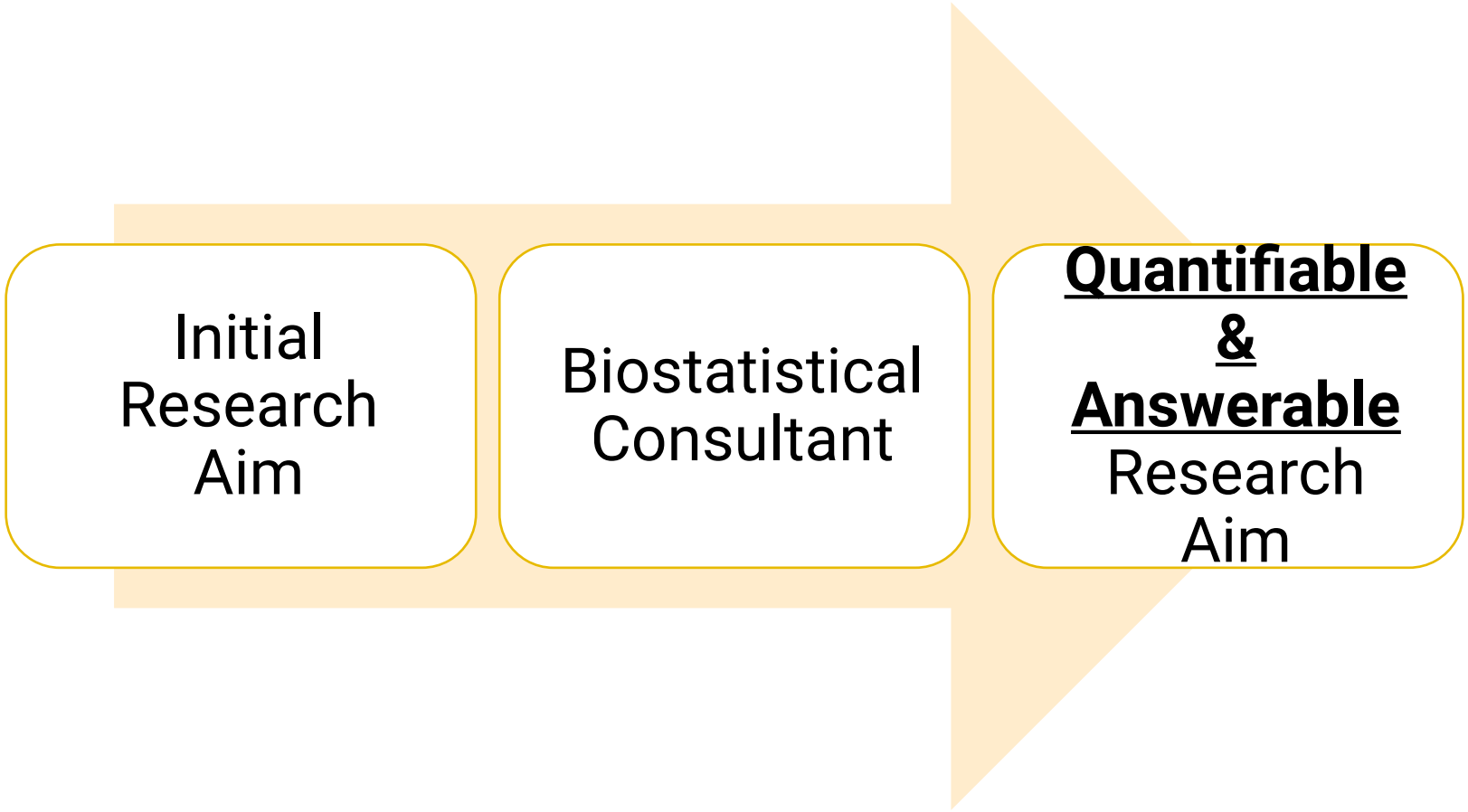
Research Aims: From Idea to Reality

- "Helps patients" compared to what?
  - Standard of care is treatment B, which many patients receive.
- How do you measure if a patient is "helped?"
  - If fewer patients progress or die from their disease.
- What other treatments could a patient have gotten before A?
  - Many patients can get radiation, which is shown to decrease the risk of progression. A majority of patients in group A get radiation.
- What patient or disease characteristics are also indicators of patient improvement?
  - Early-stage patients have a better chance of improvement, as well as younger patients. Due to the social determinants of health, patients from lower income or rural counties may have worse outcomes.

Research Aims: From Idea to Reality

- **Old**: “I want to collect all patients at our institution with disease X who received treatment A, to see if treatment A helps patients.”
- **New**: “In patients diagnosed with disease X at our institution between the years 2010 and 2019, determine the effect of novel treatment A on progression-free survival (PFS) compared to standard of care B. Account for patient and clinical characteristics that may confound the relationship between treatment and PFS. Investigate the interaction between treatment A and radiation.”

Research Aims: From Idea to Reality



Hallmarks of a Successful Project

- Communication early and often
- Clean data
- Defined set of variables and outcomes
- Deference to biostatistical expertise
- Ample time before submission deadlines (minimum of 2 weeks, prefer 4 weeks).

What Tools Do You Use?

- The Biostatistics Core maintains several SAS macros and R functions to help with basic analyses.
- About 50/50 split between coding languages
- R for data cleaning and data management
- SAS for data analysis and report creation
- I find SAS unintuitive for data importing and formatting
- I find RMarkdown, while sleek and reproducible, to be inferior to generating .rtf reports for most investigators who expect results in MS Word format
- Some R packages provide more niche analytic frameworks that SAS does not have
- Need to be adaptable to both settings!

What Statistics Do You Use?

- Much clinical cancer research is concerned with long- and short-term survival and monitoring of tumor characteristics
- **Survival methods**: Kaplan-Meier, Cox regression, time-dependent covariates, competing risks, recurrent events
- **Mixed effects methods**: Repeated measures for tumor growth curves, multiple tumors within a single patient, treated v. untreated sites in the body
- **Machine learning methods**: Radiomic features to predict patient outcomes: feature reduction & model selection
- **Standard methods**: Linear/logistic regression, chi-square, Fisher's exact, t-test, ROCAUC. I still check diagnostic plots!

Insights and Lessons from Academic Clinical Research

Cancer (and Clinical) Research is Diverse

Diversity in subject matter

- Wide array of departments I work with
  - Radiology
  - Internal Medicine
  - Dermatology
  - Urology
  - Microbiology & Immunology
  - Hematology
  - Surgery
- Opportunity to form working relationships with specific departments
  - Larger amount of dermatology projects this past year
  - Other staff include heavier volumes of urology and radiology projects

Cancer (and Clinical) Research is Diverse

Diversity in data and analysis

- Disparities in melanoma outcomes between rural and urban Iowans
- Tumor growth curves and cell activity in lab mice
- Racial disparities in genomic testing for pancreatic cancer patients
- Effect of radical bladder surgery and chemotherapy on disease recurrence
- Machine-learning to predict patient outcomes from PET/CT images
- Treatment efficacy across protein expressions and genetic mutations
- If it is cancer-related, we analyze it!

Cancer Research is Cyclical

- Abstract/manuscript submission deadlines, conferences, and grant submission windows lead to a highly cyclical workload
- Light months in early winter, early spring, and mid summer lead to more time for training and programming
- Heavy months require lots of ability to work independently and manage ~10 ongoing projects at once.
- Batch-drafting of emails leads to greater time for uninterrupted statistical analysis and maintain forward momentum
- Revolving door of projects at different stages

Statistics is Storytelling

- Analysis is only half the job. Constructing an interesting story from data leads to greater interest and uptake.
- Each row in a data set is a person's entire experience with their disease, and many times, that experience has ended prematurely.
- The "bio" in "biostatistics" is important not to forget about.
- Bio = life, a living person whose story and data should be stewarded with the utmost care, to inform clinical research conclusions.

Other Insights

- “I don’t know,” is an acceptable answer
- If working from home, increased responsibility to learn the job and stay up to date on project needs
- Clear and effective communication is your most important tool in your statistical toolbox
- Adaptability to programming, statistical methodology, changing study aims, disease subspeciality subject matter is key
- Always say yes to new opportunities. You’ll keep those skills forever!

Thank You!