

## Yuyu(Ruby) Chen

(714) 831-9627 | [yc4178@nyu.edu](mailto:yc4178@nyu.edu)

<https://rubychan299.netlify.app/>

### EDUCATION

---

- 2019 – 2020      **M.S. Biostatistics** (GPA: 3.98/4.00)  
New York University, New York, NY  
Courses: Regression, Missing Data, Survival Analysis, Bayesian, Statistical Inference, Epidemiology, Survey Methods, Psychometric Measurements
- 2015 - 2019      **B.A., Biochemistry** (Math GPA: 3.56/4.00)  
**Minor, Mathematics**  
Occidental College, Los Angeles, CA  
Courses: Calculus (I, II), Ordinary Differential Equations, Fundamentals of Computer Science, Bioinformatics
- 2017              **Study Abroad** (GPA: 85/100; Top 5%)  
Jinan University, Guangzhou, China  
Courses: Statistics, Linear Algebra, Multivariable Calculus

### RESEARCH EXPERIENCE

---

- 2019 – Present      **Measurement, Learning & Evaluation Lab (MLE), New York University**, New York, NY  
*Graduate Research Assistant*  
PI: Melody Goodman, Ph.D.  
  - Managed the data-literacy training survey data and performed analysis on pre/post-training bivariate and multivariate analysis on the association of student performance among college students and high school students the effect of remote learning.
  - Validated the research-literacy and research-knowledge evaluation scale by assessing reliability, validity, factor analysis and principal component analysis.
  - Implemented complex survey design on COVID-19 attitude, knowledge and behavior among younger population with scale development.
  - Construct the frontend and design the UI/UX of MLE lab website.
- 2019 – Present      **Center for Drug Use and HIV/HCV Research, New York University**, New York, NY  
*Biostatistician*  
PI: Melody Goodman, Ph.D.& Alex Bennett, Ph.D.& Luther Elliot, Ph.D.  
  - Directed the Opioid Overdose Risk Management Research clinical trial data management and analysis.
  - Developed Opioid Overdose Risk Behavior Score measurement scale using factor analysis and latent class analysis.
  - Conducted linear-mixed effects model on longitudinal data to determine the association between demographic, psychometric status and opioid overdose behavior.
  - Assessed the Poisson regression predictors of Opioid Overdose Risk Behavior in relation with Naloxone use, psychometric status and biophysical factors.
- 2019 – Present      **Population Health Biostatistics Division, NYU Langone Health**, New York, NY  
*Biostatistics Research Assistant*  
PI: Eva. Petkova, Ph.D. & Keith Goldfeld, Ph.D.

- Collaborated in an innovative Phase II/Phase III clinical trial study design and statistical analysis plan preparation for early pain investigation using the Bayesian Adaptive Platform trial approach.
- Implemented an “N of 1” multi-crossover study design simulation R function using simstudy to perform power analysis and sample size calculation for Centrexion’s asset.
- Developed simulations for continuous outcome in Bayesian Adaptive basket trial study design in Rstan interface to prove the efficacy.
- Performed Bayesian prior prediction simulations with random effect ordinal logistic model for COVID-19 convalescent plasma pooling project on High Performance Computing Cluster using SlurmR.

2020 – Present

**New York City Department of Health and Mental Hygiene**, New York, NY

*Naloxone Research Intern*

Supervisor: Alexandra Kingsepp

- Managed large raw Naloxone Recipients records database and queried data from 10,000 recipients using MS Access and MS SQL server.
- Completed cleaning and matching for involved Naloxone monthly distribution reports among 1000+ Opioid Prevention Program sites using R and MySQL.
- Automated pharmacy indicator, pharmacy standard report queries and the ad-hoc queries.
- Optimized the Naloxone Respondent data extract cleaning and SQL server database management processes from manual to automated program using SAS.
- Created more than 10 advanced and interactive Tableau dashboards with live connections to help Technical Assistance team monitor Opioid Overdose Prevention Programs survey responses.
- Analyzed the survey result of naloxone distribution data extract and conducted comprehensive descriptive statistics for program improvement using SQL.

2017 – 2019

**The Spain Lab, Department of Chemistry, Occidental College**, Los Angeles, CA

*Research Assistant*

Supervisor: Eileen Spain, Ph.D.

- Led bacterial incubation and sample preparation, manipulated Atomic Force Microscope to analyze bacterial adhesion.
- Analyzed force curve of bacterial adhesion using Argyle Light which further understands the mechanism and applies anti-bacterial system in medical instruments surfaces which minimized the contamination.
- Cooperated across disciplines and institutions to assist with verifying findings, sharing resources and to troubleshoot.

## CONSULTING PROJECT EXPERIENCE

Fall 2020

**The association between blood-related biomarkers and time-to-event outcome for hospitalized COVID-19 (In Progress)**

- Assessed the time to death of hospitalized COVID-19 patients Electronic Health Records (EHR) with competing risk analysis.
- Fitted biomarkers with non-linear relationship in Partial-linear single-index Cox (PLSI-Cox) regression model in association to time to death and comorbidity.

Summer 2020

**Machine Learning Loan Status/Charge-off Prediction**

- Imputed MAR and MCAR missing using mean, hotdecking and regression imputation.
- Compared machine learning models including Random Forrest, xgboost, RUS Adaboost, KNN, ANN and CNN to predict the loan status and charge-off determinations, with the best model achieving a 92% Accuracy.

Spring 2020	<b>Respondent Driven Sampling Population Estimate</b> <ul style="list-style-type: none"> <li>• Estimated the posterior probability of teenager heroin users based on survey data to create a Bayesian prior.</li> <li>• Obtained the population size estimate of the hidden population using RDS Analyst interface.</li> </ul>
Spring 2020	<b>Cessation of Heroin: A Neighborhood Grounded Exploration</b> <ul style="list-style-type: none"> <li>• Operated matching and data cleaning for a two-arm cohort study and conducted bivariate and multivariate analysis to study the heroin cessation behavior in factors including sexual transmissions.</li> </ul>

## MANUSCRIPT IN PROGRESS

---

**Chen, Y., & Goodman, M. S.** *VALIDATION OF RESEARCH KNOWLEDGE AND LITERACY EVALUATION METRICS*. Manuscript in progress.

## CONFERENCE PRESENTATIONS

---

**Chen, Y., & Spain, E. M.** (2019). *MECHANISMS OF ADHESION OF A BACTERIAL PREDATOR TO SURFACE WITH QUANTITATIVE FORCE MEASUREMENTS*. Baltimore, MD: Biophysical Society 63<sup>rd</sup> Annual Meeting 2019.

**Chen, Y.** (2018). *ADHESION OF A BACTERIAL PREDATOR TO SURFACE WITH QUANTITATIVE FORCE MEASUREMENTS*. Pasadena, CA: Southern California Conference for Undergraduate Research.

**Chen, Y.** (2018). *ADHESION OF A BACTERIAL PREDATOR TO SURFACE WITH QUANTITATIVE FORCE MEASUREMENTS*. Pasadena, CA: Undergraduate Research Center Summer Research Conference.

## TEACHING AND SERVICES

---

2020	<b>Lab Coordinator</b> , Measurement, Learning & Evaluation Lab, New York University
2020	<b>Course Assistant</b> , Social Network Analysis in R, New York University
2020	<b>Course Assistant</b> , Data-Literacy Training Summer Program
2018	<b>Laboratory Teaching Assistant</b> , General Chemistry, Occidental College
2017	<b>Recruitment Director</b> , Asian American Tutorial Project, Occidental College
2016	<b>International Ambassador</b> , Occidental College

## AWARDS AND HONORS

---

2019-2021	Master Merit-based Scholarship, New York University
2017-2019	Summer Research Program Fellowship, Occidental College
2018-2019	Academic Research Project Fellowship, Occidental College
2019	Academic Research Conference Travel Fund, Occidental College

## SKILLS

---

**Technical:** R, SAS, SQL, Stata, Tableau, Java, Python, Git, Latex, Spark, Hadoop

**Libraries:** SciKit, Pandas, matplotlib, ggplot2, MICE, dplyr, randomForest, KernLab, nnet, simstudy, igraph, Stan, Slurm

**Survey:** Qualtrics, Redcap