Naim U. Rashid

PERSONAL

University of North Carolina-Chapel Hill
Department of Biostatistics
3104-E McGavran-Greenberg Hall, CB 7420

919-843-3311
naim@unc.edu

Lineberger Comprehensive Cancer Center
LCCC 20-020, CB 7295
Chapel Hill, North Carolina 27599-7420

919-966-8150
https://naimurashid.github.io/

EDUCATION

2006 - 2013	PhD, Biostatistics University of North Carolina at Chapel Hill, Chapel Hill, NC
2002 - 2006	B.S. Biology with Pharmacology Concentration, Mathematics Minor Duke University, Durham, NC

PROFESSIONAL EXPERIENCE

07/2021 - Current	**Associate Professor**, Department of Biostatistics and the Lineberger Comprehensive Cancer Center, UNC-Chapel Hill, Chapel Hill, NC
01/2015 - 06/2021	**Assistant Professor**, Department of Biostatistics and the Lineberger Comprehensive Cancer Center, UNC-Chapel Hill, Chapel Hill, NC
08/2013 - 12/2014	**Postdoctoral Research Fellow**, Department of Biostatistics, Harvard School of Public Health, and Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute, Boston, MA

HONORS AND AWARDS

2023	Teaching Innovation Award, Gillings School of Global Public Health
2023	James E. Grizzle Distinguished Alumnus Award, Department of Biostatistics, Gillings School of Global Public Health
2021	Delta Omega Faculty Award, Gillings School of Global Public Health
2017	IBM and R.J. Reynolds Junior Faculty Development Award, UNC-CH
2013	Barry H. Margolin Dissertation Award for best doctoral dissertation completed in 2013
2006-2011	Genomics and Cancer Training Grant recipient

PROFESSIONAL MEMBERSHIPS

American Statistical Association

Eastern North American Region of the International Biometric Society

BIBLIOGRAPHY

Methodology (* indicates student)

1. H. M. Heiling*, **N.U. Rashid**, Q. Li, X. L. Peng, J. J. Yeh, and J. G. Ibrahim. Efficient computation of high-dimensional penalized piecewise constant hazard random effects survival models. *Journal of Computational and Graphical Statistics, Submitted*, 2023

2. H. M. Heiling*, **N.U. Rashid**, Q. Li, X. L. Peng, J. J. Yeh, and J. G. Ibrahim. Efficient computation of high-dimensional penalized generalized linear mixed models by latent factor modeling of the random effects. *Biometrics, Revision Invited*, 2023

3. D. K. Lim*, **N.U. Rashid**, J. B. Oliva, and J. G. Ibrahim. Deeply-learned generalized linear models with missing data. *Journal of Computational and Graphical Statistics, Revision Invited*, 2023

4. H. Heiling*, **N.U. Rashid**, Q. Li, and J. G. Ibrahim. High dimensional penalized generalized linear mixed models: The glmmpen r package. *The R Journal, Under 2nd Round Review*, 2023

5. D. K. Lim*, **N.U. Rashid**, J. B. Oliva, and J. G. Ibrahim. Handling non-ignorably missing features in electronic health records data using importance-weighted autoencoders. *Biometrics, Revision Invited*, 2023

6. A. M. Young*, S. Van Buren*, and **N.U. Rashid**. Differential transcript usage analysis incorporating quantification uncertainty via compositional measurement error regression modeling. *Biostatistics, In Press*, 2023

7. J. Leary, Y. Xu, A. B. Morrison, C. Jin, E. Shen, P. C. Kuhlers, Y. Su, **N.U. Rashid**, J. J. Yeh, and P. Xianlu. Sub-cluster identification through semi-supervised optimization of rare-cell silhouettes (scissors) in single-cell rna-sequencing. *Bioinformatics*, 39(8):btad449, 2023

8. H. M. Heiling*, D. R. Wilson, **N.U. Rashid**, W. Sun, and J. G. Ibrahim. Estimating cell type composition using isoform expression one gene at a time. *Biometrics*, 79(2):854–865, 2023

9. P. L. Baldoni*, **N.U. Rashid**, and J. G. Ibrahim. Efficient detection and classification of epigenomic changes under multiple conditions. *Biometrics*, 78(3):1141–1154, 2022

10. S. Van Buren*, H. Sarkar, A. Srivastava, **N.U. Rashid**, R. Patro, and M. I. Love. Compression of quantification uncertainty for scrna-seq counts. *Bioinformatics*, 37(12):1699–1707, 2021

11. D. K. Lim*, **N.U. Rashid**, and J. G. Ibrahim. Model-based feature selection and clustering of rna-seq data for unsupervised subtype discovery. *The annals of applied statistics*, 15(1):481, 2021

12. **N.U. Rashid**, D. J. Luckett, J. Chen, M. T. Lawson, L. Wang, Y. Zhang, E. B. Laber, Y. Liu, J. J. Yeh, D. Zeng, et al. High-dimensional precision medicine from patient-derived xenografts. *Journal of the American Statistical Association*, pages 1–15, 2020

13. **N.U. Rashid**, X. L. Peng, C. Jin, R. A. Moffitt, K. E. Volmar, B. A. Belt, R. Z. Panni, T. M. Nywening, S. G. Herrera, K. J. Moore, et al. Purity independent subtyping of tumors (purist), a clinically robust, single-sample classifier for tumor subtyping in pancreatic cancer. *Clinical Cancer Research*, 26(1):82–92, 2020

14. **N.U. Rashid**, Q. Li, J. J. Yeh, and J. G. Ibrahim. Modeling between-study heterogeneity for improved reproducibility in gene signature selection and clinical prediction. *Journal of the American Statistical Association*, 115(531):1125–1138, 2020

15. P. L. Baldoni*, **N.U. Rashid**, and J. G. Ibrahim. Improved detection of epigenomic marks with mixed-effects hidden markov models. *Biometrics*, 75(4):1401–1413, 2019

16. **N.U. Rashid**, W. Sun, and J. G. Ibrahim. A statistical model to assess (allele-specific) associations between gene expression and epigenetic features using sequencing data. *Annals of Applied Statistics*, 10(4):2254, 2016

17. **N.U. Rashid**, A. S. Sperling, N. Bolli, D. C. Wedge, P. Van Loo, Y.-T. Tai, M. A. Shammas, M. Fulciniti, M. K. Samur, P. G. Richardson, et al. Differential and limited expression of mutant alleles in multiple myeloma. *Blood*, 124(20):3110–3117, 2014

18. **N.U. Rashid**, W. Sun, and J. G. Ibrahim. Some statistical strategies for dae-seq data analysis: variable selection and modeling dependencies among observations. *Journal of the American Statistical Association*, 109(505):78–94, 2014

19. **N.U. Rashid**, P. G. Giresi, J. G. Ibrahim, W. Sun, and J. D. Lieb. Zinba integrates local covariates with dna-seq data to identify broad and narrow regions of enrichment, even within amplified genomic regions. *Genome biology*, 12(7):R67, 2011

Collaborative - Genomics and Cancer

20. L. A. Torre-Healy, R. R. Kawalerski, K. Oh, L. Chrastecka, X. L. Peng, A. J. Aguirre, **N.U. Rashid**, J. J. Yeh, and R. A. Moffitt. Open-source curation of a pancreatic ductal adenocarcinoma gene expression analysis platform (pdacr) supports a two-subtype model. *Communications Biology*, 6(1):163, 2023

21. A. Fernandez-Martinez, T. Pascual, B. Singh, P. Nuciforo, **N.U. Rashid**, K. V. Ballman, J. D. Campbell, K. A. Hoadley, P. A. Spears, L. Pare, et al. Prognostic and predictive value of immune-related gene expression signatures vs tumor-infiltrating lymphocytes in early-stage erbb2/her2-positive breast cancer: A correlative analysis of the calgb 40601 and pamela trials. *JAMA oncology*, 2023

22. D. McGrail, P. Pilié, **N.U. Rashid**, L. Voorwerk, M. Slagter, M. Kok, E. Jonasch, M. Khasraw, A. Heimberger, N. Ueno, et al. Validation of cancer-type-dependent benefit from immune checkpoint blockade in tmb-h tumors identified by the foundationone cdx assay. *Annals of Oncology*, 33(11):1204–1206, 2022

23. M. V. Huynh, G. A. Hobbs, A. Schaefer, M. Pierobon, L. M. Carey, J. N. Diehl, J. M. DeLiberty, R. D. Thurman, A. R. Cooke, C. M. Goodwin, et al. Functional and biological heterogeneity of krasq61 mutations. *Science Signaling*, 15(746):eabn2694, 2022

24. B. Mirlekar, Y. Wang, S. Li, M. Zhou, S. Entwistle, T. De Buysscher, A. Morrison, G. Herrera, C. Harris, B. G. Vincent, et al. Balance between immunoregulatory b cells and plasma cells drives pancreatic tumor immunity. *Cell Reports Medicine*, 3(9):100744, 2022

25. F. Innocenti, A. Yazdani, **N.U. Rashid**, X. Qu, F.-S. Ou, S. Van Buren, M. M. Bertagnolli, O. Kabbarah, C. D. Blanke, A. P. Venook, H.-J. Lenz, and B. G. Vincent. Tumor Immunogenomic Features Determine Outcomes in Patients with Metastatic Colorectal Cancer Treated with Standard-of-Care Combinations of Bevacizumab and Cetuximab. *Clinical Cancer Research*, 28(8):1690–1700, 04 2022

26. J. N. Diehl, J. E. Klomp, K. R. Snare, P. S. Hibshman, D. R. Blake, Z. D. Kaiser, T. S. Gilbert, E. Baldelli, M. Pierobon, B. Papke, et al. The kras-regulated kinome identifies wee1 and erk coinhibition as a potential therapeutic strategy in kras-mutant pancreatic cancer. *Journal of Biological Chemistry*, 297(5), 2021

27. D. Zeitouni, M. P. Catalino, J. Wise, S. McCabe, K. Pietrosimone, **N.U. Rashid**, and S. Khagi. Clinical application of next-generation sequencing in recurrent glioblastoma. *Onco*, 1(1):38–48, 2021

28. D. McGrail, P. Pilié, **N.U. Rashid**, L. Voorwerk, M. Slagter, M. Kok, E. Jonasch, M. Khasraw, A. Heimberger, N. Ueno, et al. Reply to:"real-world prevalence across 159,872 patients with cancer supports the clinical utility of tmb-h to define metastatic solid tumors for treatment with pembrolizumab." by d. fabrizio et al. *Annals of Oncology*, 2021

29. S. P. Angus, T. J. Stuhlmiller, G. Mehta, S. M. Bevill, D. R. Goulet, J. F. Olivares-Quintero, M. P. East, M. Tanioka, J. S. Zawistowski, D. Singh, et al. Foxa1 and adaptive response determinants to her2 targeted therapy in tbcrc 036. *NPJ breast cancer*, 7(1):1–15, 2021

30. D. McGrail, P. Pilié, **N.U. Rashid**, L. Voorwerk, M. Slagter, M. Kok, E. Jonasch, M. Khasraw, A. Heimberger, B. Lim, et al. High tumor mutation burden fails to predict immune checkpoint blockade response across all cancer types. *Annals of Oncology*, 32(5):661–672, 2021

31. M. B. Lipner, X. L. Peng, C. Jin, Y. Xu, Y. Gao, M. P. East, **N.U. Rashid**, R. A. Moffitt, S. G. Herrera-Loeza, A. B. Morrison, B. T. Golitz, C. Vaziri, L. M. Graves, G. L. Johnson, and J. J. Yeh. Irreversible JNK1-JUN inhibition by JNK-IN-8 sensitizes pancreatic cancer to 5-FU/FOLFOX chemotherapy. *JCI Insight*, 5(8), apr 2020

32. B. Mirlekar, D. Michaud, S. J. Lee, N. P. Kren, C. Harris, K. Greene, E. C. Goldman, G. P. Gupta, R. C. Fields, W. G. Hawkins, D. G. DeNardo, **N.U. Rashid**, J. J. Yeh, A. J. McRee, B. G. Vincent, D. A. Vignali, and Y. Pylayeva-Gupta. B cell–derived il35 drives stat3-dependent cd8+ t-cell exclusion in pancreatic cancer. *Cancer immunology research*, 8(3):292–308, 2020

33. K. D. Fagan-Solis, D. A. Simpson, R. J. Kumar, L. G. Martelotto, L. E. Mose, **N.U. Rashid**, A. Y. Ho, S. N. Powell, Y. H. Wen, J. S. Parker, et al. A p53-independent dna damage response suppresses oncogenic proliferation and genome instability. *Cell Reports*, 30(5):1385–1399, 2020

34. W. Feng, D. A. Simpson, J. Carvajal-Garcia, B. A. Price, R. J. Kumar, L. E. Mose, R. D. Wood, **N.U. Rashid**, J. E. Purvis, J. S. Parker, et al. Genetic determinants of cellular addiction to dna polymerase theta. *Nature communications*, 10(1):1–13, 2019

35. J. A. Wrobel, L. Xie, L. Wang, C. Liu, **N.U. Rashid**, K. K. Gallagher, Y. Xiong, K. D. Konze, J. Jin, M. L. Gatza, et al. Multi-omic dissection of oncogenically active epiproteomes identifies drivers of proliferative and invasive breast tumors. *iScience*, 17:359–378, 2019

36. S. M. Bevill, J. F. Olivares-Quintero, N. Sciaky, B. T. Golitz, D. Singh, A. S. Beltran, **N.U. Rashid**, T. J. Stuhlmiller, A. Hale, N. J. Moorman, et al. Gsk2801, a baz2/brd9 bromodomain inhibitor, synergizes with bet inhibitors to induce apoptosis in triple-negative breast cancer. *Molecular Cancer Research*, 17(7):1503–1518, 2019

37. A. V. Vaseva, D. R. Blake, T. S. Gilbert, S. Ng, G. Hostetter, S. H. Azam, I. Ozkan-Dagliyan, P. Gautam, K. L. Bryant, K. H. Pearce, et al. Kras suppression-induced degradation of myc is antagonized by a mek5-erk5 compensatory mechanism. *Cancer cell*, 34(5):807–822, 2018

38. Y. Yang, O. Adebali, G. Wu, C. P. Selby, Y.-Y. Chiou, **N.U. Rashid**, J. Hu, J. B. Hogenesch, and A. Sancar. Cisplatin-dna adduct repair of transcribed genes is controlled by two circadian programs in mouse tissues. *Proceedings of the National Academy of Sciences*, 115(21):E4777–E4785, 2018

39. R. J. Torphy, Z. Wang, A. True-Yasaki, K. E. Volmar, **N.U. Rashid**, B. Yeh, J. S. Johansen, M. A. Hollingsworth, J. J. Yeh, and E. A. Collisson. Stromal content is correlated with tissue site, contrast retention, and survival in pancreatic adenocarcinoma. *JCO precision oncology*, 2:1–12, 2018

40. A. E. Van Swearingen, M. J. Sambade, M. B. Siegel, S. Sud, R. S. McNeill, S. M. Bevill, X. Chen, R. E. Bash, L. Mounsey, B. T. Golitz, et al. Combined kinase inhibitors of mek1/2 and either pi3k or pdgfr are efficacious in intracranial triple-negative breast cancer. *Neuro-oncology*, 19(11):1481–1493, 2017

41. K. C. Arend, E. M. Lenarcic, H. A. Vincent, **N.U. Rashid**, E. Lazear, I. M. McDonald, T. S. Gilbert, M. P. East, L. E. Herring, G. L. Johnson, et al. Kinome profiling identifies druggable targets for novel human cytomegalovirus (hcmv) antivirals. *Molecular & Cellular Proteomics*, 16(4 suppl 1):S263–S276, 2017

42. J. S. Zawistowski, S. M. Bevill, D. R. Goulet, T. J. Stuhlmiller, A. S. Beltran, J. F. Olivares-Quintero, D. Singh, N. Sciaky, J. S. Parker, **N.U. Rashid**, et al. Enhancer remodeling during adaptive bypass to mek inhibition is attenuated by pharmacologic targeting of the p-tefb complex. *Cancer discovery*, 7(3):302–321, 2017

43. D. R. Roque, L. Makowski, T.-H. Chen, **N.U. Rashid**, D. N. Hayes, and V. Bae-Jump. Association between differential gene expression and body mass index among endometrial cancers from the cancer genome atlas project. *Gynecologic oncology*, 142(2):317–322, 2016

44. Y.-Y. Chiou, Y. Yang, **N.U. Rashid**, R. Ye, C. P. Selby, and A. Sancar. Mammalian period represses and de-represses transcription by displacing clock–bmal1 from promoters in a cryptochrome-dependent manner. *Proceedings of the National Academy of Sciences*, 113(41):E6072–E6079, 2016

45. A. R. Johnson, Y. Qin, A. J. Cozzo, A. J. Freemerman, M. J. Huang, L. Zhao, B. P. Sampey, J. J. Milner, M. A. Beck, B. Damania, et al. Metabolic reprogramming through fatty acid transport protein 1 (fatp1) regulates macrophage inflammatory potential and adipose inflammation. *Molecular metabolism*, 5(7):506–526, 2016

46. R. H. Prabhala, M. Fulciniti, D. Pelluru, **N.U. Rashid**, A. Nigroiu, P. Nanjappa, C. Pai, S. Lee, N. S. Prabhala, R. L. Bandi, et al. Targeting il-17a in multiple myeloma: a potential novel therapeutic approach in myeloma. *Leukemia*, 30(2):379, 2016

47. R. A. Moffitt, R. Marayati, E. L. Flate, K. E. Volmar, S. G. H. Loeza, K. A. Hoadley, **N.U. Rashid**, L. A. Williams, S. C. Eaton, A. H. Chung, et al. Virtual microdissection identifies distinct tumor-and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nature genetics*, 47(10):1168, 2015

48. M. Shapiro, **N.U. Rashid**, E. E. Whang, V. A. Boosalis, Q. Huang, C. Yoon, M. S. Saund, and J. S. Gold. Trends and predictors of resection of the primary tumor for patients with stage iv colorectal cancer. *Journal of surgical oncology*, 111(7):911–916, 2015

49. M. Shapiro, **N.U. Rashid**, Q. Huang, S. L. Galper, V. A. Boosalis, E. E. Whang, and J. S. Gold. Radiation therapy for unresectable pancreatic adenocarcinoma: population-based trends in utilization and survival rates in the united states. *JAMA surgery*, 150(3):274–277, 2015

50. N. Bolli, H. Avet-Loiseau, D. C. Wedge, P. Van Loo, L. B. Alexandrov, I. Martincorena, K. J. Dawson, F. Iorio, S. Nik-Zainal, G. R. Bignell, et al. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nature communications*, 5:2997, 2014

51. B. Bernstein, E. Birney, I. Dunham, E. Green, C. Gunter, M. Snyder, et al. An integrated encyclopedia of dna elements in the human genome. *Nature*, 489(7414):57, 2012

Collaborative - Other

52. F. A. Oladosu, M. S. Conrad, S. C. O'Buckley, **N.U. Rashid**, G. D. Slade, and A. G. Nackley. Mu opioid splice variant mor-1k contributes to the development of opioid-induced hyperalgesia. *PloS one*, 10(8):e0135711, 2015

53. I. Belfer, S. K. Segall, W. R. Lariviere, S. B. Smith, F. Dai, G. D. Slade, **N.U. Rashid**, J. S. Mogil, C. M. Campbell, R. R. Edwards, et al. Pain modality-and sex-specific effects of comt genetic functional variants. *PAIN*, 154(8):1368–1376, 2013

54. D. Tsao, J. S. Wieskopf, **N.U. Rashid**, R. E. Sorge, R. L. Redler, S. K. Segall, J. S. Mogil, W. Maixner, Q. Zheng, D. Fang, et al. Serotonin-induced hypersensitivity via inhibition of catechol o-methyltransferase activity. *Molecular Pain*, 8(1):25, 2012

55. G. D. Slade, M. S. Conrad, L. Diatchenko, **N.U. Rashid**, S. Zhong, S. Smith, J. Rhodes, A. Medvedev, S. Makarov, W. Maixner, et al. Cytokine biomarkers and chronic pain: association of genes, transcription, and circulating proteins with temporomandibular disorders and widespread palpation tenderness. *Pain*, 152(12):2802–2812, 2011

PATENTS

1. R. Moffitt, J. J. Yeh, and N. U. Rashid. Methods and compositions for prognostic and/or diagnostic subtyping of pancreatic cancer, Apr. 28 2022. US Patent App. 17/336,600

2. R. Moffitt, J. J. Yeh, and **N.U. Rashid**. Gene-expression based subtyping of pancreatic ductal adenocarcinoma, July 6 2021. US Patent 11,053,550

ORAL PRESENTATIONS

Invited

2023	A robust and clinically applicable classifier for PDAC stromal subtyping. PDAC Stromal Reprogramming Consortium Steering Committee Annual Meeting, Ann Arbor, MI.
2023	Robust and replicable supervised and unsupervised learning methods for cancer precision medicine. IISA Annual Conference, Golden, CO.
2022	Cancer Data Science: From Code to Clinic. Seminar, Carolina Data Science Now, University of North Carolina, Chapel Hill, NC.

Invited

2022	Addressing the Replicability and Generalizability of Genomic Prediction Models. Seminar, Hillman Cancer Center, University of Pittsburgh, Pittsburgh, PA.
2022	Missing Data Methods for Supervised and Unsupervised Deep Learning Tasks. Seminar, Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD.
2021	Addressing the Replicability and Generalizability of Genomic Prediction Models. BIS Seminar, Department of Biostatistics, Yale School of Public Health, New Haven, CT.
2021	Replicability and missing data in deep learning and clinical prediction. AAAS invited session, JSM, Seattle, WA.
2021	Cancer Biostatistics Research at UNC. Green Level High School, Cary, NC.
2021	Missing Data Methods for Supervised and Unsupervised Deep Learning Tasks. Seminar, Department of Biostatistics, Indiana Univeristy School of Medicine, Indianapolis, IN.
2020	Research Developments and Opportunities in AI and Precision Health at UNC Biostatistics. AI and Health Meeting, Chapel Hill, NC.
2019	PurIST: a clinically robust single sample classifier for tumor subtyping in pancreatic cancer. A Consensus Workshop for Pancreatic Ductal Adenocarcinoma Taxonomy, MSKCC, NY, NY.
2019	Modeling Between-Study Heterogeneity for Improved Reproducibility in Gene Signature Selection and Clinical Prediction. ENAR, Philadelphia, PA.
2018	Modeling Between-Study Heterogeneity for Improved Reproducibility in Gene Signature Selection and Clinical Prediction. AISC, Greensboro, NC.
2014	How Its Done Series: Next Generation Sequencing Pipelines. Dana Farber Cancer Institute and Harvard School of Public Health, Boston, MA. Seminar
2013	Some Statistical Strategies for DAE-seq Data Analysis: variable selection and Modeling Dependencies among Observations. Dana Farber Cancer Institute and Harvard School of Public Health, Boston, MA.
2010	Addressing emerging challenges of ChIP-seq data analysis. North Carolina Biotechnology Center, Durham, NC.

Contributed

2021	Analyzing the Data to Ensure Reproducibility. Joint T32 Rigor and Reproducibility Seminar Series, Lineberger Comprehensive Cancer Center, Chapel Hill, NC.
2017	High Dimensional Precision Medicine in Patient-Derived Xenografts. JSM, Baltimore, MD.
2017	Addressing Between-Study Heterogeneity for Improved Reproducibility in Gene Signature Selection and Clinical Prediction. JSM, Baltimore, MD.
2016	Robust Approaches for the Analysis of High-Throughput Proteomic Data. JSM, Chicago, IL.
2015	Fast and flexible determination of differential alternative splicing from RNA-seq data. JSM, Seattle, WA.
2014	Efficient and scalable approaches for the detection of differential alternative splicing from RNA-seq data. DFCI-BCB Genomics Get Together Seminar, Boston, MA.
2014	Alternative Splicing Is a Frequent Event and Impacts Clinical Outcome in Myeloma: A Large RNA-Seq Data Analysis of Newly-Diagnosed Myeloma Patients. American Society for Hematology Meeting, San Francisco, CA.
2013	Applications of RNA-seq in Multiple Myeloma. VA Boston Medical Center, West Roxbury, MA.
2012	Autoregressive modelling and variable selection procedures in hidden markov models with covariates, with applications to DAE-seq data. Seminar, Biostatistics Core, Lineberger Comprehensive Cancer Center, Chapel Hill, NC.
2011	Problems in next generation sequencing analysis. Guest Lecture, Bios 782: Methods in Computational Biology, Chapel Hill, NC.
2011	Mixture regression analysis of *-Seq data. ENAR, Miami, FL.
2010	ZINBA: A unified modeling framework for the analysis and exploration of diverse ChIP-seq signal patterns. ModENCODE group meeting, Chapel Hill, NC.
2010	Next generation sequencing: analysis and inference. Seminar, Strahl-Davis-Lieb joint lab meeting, Chapel Hill, NC.
2010	Research in functional genomics. UNC Biostatistics Prospective Students Day, Chapel Hill, NC.
2006	Finite sample properties of estimators of the False Discovery Rate. ENAR, Tampa Bay, FL.

SOFTWARE

dlglm: R Package for flexible handling of non-ignorable missing data in deeply learned generalized linear models. Developed with David Lim and available at
`https://github.com/DavidKLim/dlglm`

glmmPen: R Package for simultaneous fixed and random effects selection in high dimensional generalized linear mixed models. Developed with Hillary Heiling and available at `https://github.com/hheiling/glmmPen/`

NIMIWAE: R Package for flexible handling and imputation of non-ignorable missing data patterns using Deep Learning Variational Autoencoders. Developed with David Lim and available at `https://github.com/DavidKLim/NIMIWAE`

epigraHMM: Bioconductor package for multi-sample consensus and differential enrichment pattern detection from ChIP-seq, ATAC-seq, and related data types. Developed with Pedro Baldoni and available at `http://bioconductor.org/packages/release/bioc/html/epigraHMM.html`

FSCseq: Computational method to simulatenously detect latent clusters and cluster-discriminatory genes from RNA-seq data. Developed with David Lim and available at `https://github.com/DavidKLim/FSCseq`

mixNBHMM: A highly efficient and flexible algorithm for calling differential peaks across in multi-sample, multi-condition experimental settings for ChIP-seq, ATAC-seq, DNase-seq, and similar data. Can also be applied to multiple types of ChIP-seq experiments from the same condition to determine combinatorial patterns of interactions between different epigenomic processes across the genome. Developed with Pedro Baldoni and available at `https://github.com/plbaldoni/mixNBHMM`

ZIMHMM: An HMM-based algorithm for calling broad consensus regions of enrichment across multiple technical ChIP-seq replicates. Developed with Pedro Baldoni and available at `https://github.com/plbaldoni/ZIMHMM`

Zero Inflated Negative Binomial Algorithm: A comprehensive R package for the statistical detection of genomic regions enriched for NGS reads and applicable to a wide variety of NGS datasets; available at code.google.com/p/zinba

hmmcov: An R package for the analysis of DAE-seq data implementing HMM and AR-HMM based procedures for enrichment detection in epigenetic datasets. Also implements novel variable selection procedure for the efficient detection of biological factors associated with correlated genomic features; available at code.google.com/p/hmmcov

BASeG: An R package for Bivariate Aassociation studies using Sequencing data, while accounting for shared Genetic effects. Bivariate Poisson-lognormal and Bivariate Logistic-normal regression is utilized to assess the associations between gene expression and epigenetic marks from sequencing data, while explicitly modeling the effects of DNA polymorphisms in either an allele-specific or non-allele-specific manner.

TEACHING RECORD

Spring 2019-2023	Bios 735: Intro to Statistical Computing Department of Biostatistics University of North Carolina at Chapel Hill *Course Co-Developer (with Dr. Mike Love) and Instructor* *Required Course, 4 credit hours, 30 students*

Spring 2016-2018 Bios 663: Intermediate Linear Models
Department of Biostatistics
University of North Carolina at Chapel Hill
Course Developer and Instructor
Required Course, 4 credit hours, 35 students

BIOS DOCTORAL STUDENTS ADVISED

1. Pedro Baldoni (2016-2020), coadvised with Dr. Joseph Ibrahim
2. David Lim (2016-2022), coadvised with Dr. Joseph Ibrahim
3. Scott Van Buren (2017-2020), coadvised with Dr. Mike Love
4. Hillary Heiling (2018-2023 expected), coadvised with Dr. Joseph Ibrahim
5. Christina Zhou (2020-), coadvised with Dr. Joseph Ibrahim
6. Euphy Wu (2021-), coadvised with Dr. Mike Love
7. Amber Young (2021-)

BIOS MPH DATA SCIENCE STUDENTS ADVISED

1. Aarushi Jothi (2019-2021)
2. Zhitong Yu (2020-)
3. Surya Sampath (2021-)

BIOS UNDERGRADUATE STUDENTS ADVISED

1. Tianyi Liu (2019), BSPH Honors Thesis

PHD DISSERTATION COMMITTEE

1. Doug Wilson, Department of Biostatistics (2016). Advisor: Joseph Ibrahim & Wei Sun
2. Vasyl Zhabotynsky, Department of Biostatistics (2016). Advisor: Wei Sun
3. Heejoon Jo, Department of Biostatistics (2017). Advisor: Neil Hayes & Steve Marron
4. Aatish Thennavan, Oral and Craniofacial Biomedicine (2017). Advisor: Chuck Perou
5. David Pritchard, Department of Biostatistics (2017). Advisor: Yufeng Liu & Matt Psioda
6. Anqi Zhu, Department of Biostatistics (2018). Advisor: Mike Love & Joseph Ibrahim
7. Brady Nifong, Department of Biostatistics (2019). Advisor: Matt Psioda & Joseph Ibrahim
8. Jiawei Xu, Department of Biostatistics (2019). Advisor: Matt Psioda & Joseph Ibrahim
9. Sean McCabe, Department of Biostatistics (2019). Advisor: Mike Love & Danyu Lin

10. Arjun Bhattacharya, Department of Biostatistics (2019). Advisor: Mike Love & Melissa Troester
11. William Belzak, Department of Psychology (2020).
12. Laura Zhou, Department of Biostatistics (2020). Advisor: Fei Zou & Wei Sun
13. Evan Kwiatkowski, Department of Biostatistics (2021). Advisor: Matt Psioda
14. Ian Sturgill, Bioinformatics and Computational Biology Program (2022). Advisors: Katie Hoadley and Jessie Raab
15. Brooke Felsheim, Bioinformatics and Computational Biology Program (2022). Advisor: Chuck Perou
16. Mikayla Feldbauer, Bioinformatics and Computational Biology Program (2023). Advisor: Katie Hoadley

GRANTS (as PI, Co-PI, or Core Leader)

Active

1. U01 07/01/2022 – 06/30/2027 1.2 Calendar
   NIH/NCI (Co-PIs: Yeh, Rashid) $4,618,709
   Integrating tumor and stroma to understand and predict treatment response
   We propose to establish predictive models that will link phenotypic responses in tissues to underlying changes in molecular states. Models will be built on data derived from our experimental platform that will characterize transcriptomic, proteomic, and phosphoproteomic states in response to well-defined targeted treatments across a range of tissue complexity; from tumor-stroma cell line mixtures, to organoids, to tumors. Together, these efforts will provide a novel opportunity to establish predictive tools that improve our ability to discern tissue-specific vulnerabilities.

2. P50-CA257911 07/31/2022-07/31/2027 3.0 Cal
   NIH-NCI (PI: Yeh) $11,086,064
   SPORE in Pancreatic Cancer (Core C: Integrated Quantitative Sciences Core)
   The Integrated Quantitative Sciences Core provides participates in the design of all clinical trials, animal studies, and translational research proposed in the SPORE to ensure that all relevant studies are well powered, utilize appropriate statistical methods, and are properly designed to address relevant hypotheses of study aims. In this manner the Core supports the rigor and reproducibility of all results that are generated by the SPORE, which in turn have significant impacts in the fields of public health and medicine. Expert analysis of project data and clear reporting of scientific results are of similar importance for addressing scientific hypotheses and similarly have strong implications in public health. **Core Co-Leader with Michael Kosorok**.

3. Alliance-5117778 7/01/2020-6/30/2023 (NCE) 1.2 Calendar
   Alliance for Clinical Trials in Oncology Foundation (PI: Rashid) $198,651
   Machine learning methods for biomarker-driven optimal treatment selection in metastatic colorectal cancer

The objective of this research is to evaluate a machine learning approach to identify individualized treatment rules for optimal selection of treatment of patients with metastatic colorectal cancer.

Completed

4. GM105785-05 5/1/17-4/30/18 0.6 Calendar
NIGMS/Fred Hutchinson (Subcontract PI: Rashid) $45,735
Statistical Methods for RNA-seq Data Analysis
This proposal will develop statistical methods to analyze RNA-seq data, with particular focus on allele-specific expression (ASE). We develop several statistical methods to dissect the effects of different factors underlying the imbalances in ASE. Our methods and results will provide much needed tools to analyze RNA-seq data, as well as insights into the regulation of gene expression.

GRANTS (as Co-I or Biostatistician)

Active

1. R01-CA270792 12/01/2022 – 11/30/2027 0.48 Cal
NIH/NCI (PI: Earp/Pylayeva-Gupta) $2,686,355
Divergent Roles of MerTK, Tyro3, and Axl in Pancreatic Cancer and Metastasis
A major reason for the therapeutic failure in pancreatic ductal cancer is the pro-tumorigenic function of myeloid cells and fibroblasts that produce an immunosuppressive tumor microenvironment. This attribute has resulted in clinical failure of the newer modalities of cancer immunotherapy. We have identified that MerTK, Tyro3 and Axl receptor tyrosine kinases are important but distinct regulators of tumor growth and T cells responses in pancreatic cancer. Proposed research will elucidate mechanisms by which these kinases promote or impede pancreatic tumor growth and provide us with novel therapeutic targets that could be used to synergize with existing therapeutic approaches.

2. Not assigned 01/02/2023 – 01/01/2026 0.24 Cal
Lustgarten Foundation (PI: Yeh) $450,000
PROmoting CLinicAl TrIal EngageMent for Pancreatic Cancer App Study (PROCLAIM Study)
This proposal will use mHealth technology as an educational, communication, and audit and feedback tool to promote patient-initiated clinical trial discussions among Black people with pancreas cancer with their cancer care team.

3. RSG-21-103-01 01/01/2022-12/31/2025 0.12 Cal
American Cancer Society (PI: Pylayeva-Gupta) $792,000
B cells as mediators of tumor eradication in pancreatic cancer
T cells are a part of the immune system able to control tumor growth, but are ineffective in most cancer patients. Our lab has made the groundbreaking discovery that T cells entering tumors experience an extreme response to stress that rapidly shuts down their ability to generate energy and make and process proteins. Our proposal will formally prove that targeting the stress response holds the power to revive essential T cell functions and invigorate tumor control to allow multiple types of immunotherapies to work better for cancer patients.

4. P30-CA016086 12/21/2020 – 11/30/2025 3.60 Calendar
NIH/NCI (PI: Earp) $288,650.00

Cancer Center Core Support Grant: - Biostatistics
Biostatistics Core - Areas of expertise in support of future use by LCCC members include clinical trials and other forms of clinical research, computational biology, genomics, genetics, cancer epidemiology, quality of life, outcomes, and other forms of population sciences research.

5. OG22874046 06/08/2022 – 07/07/2025 0.6 Calendar
Komen Foundation (PI: Hoadley/Freedman) $160,816.00
Ancestry-related RNA Splicing and Immune Expression in Metastatic Breast Cancer
This project will explore the role of RNA splicing in metastatic breast cancer and determine if there differences by race.

6. R37-C247676 07/01/2020-06/30/2025 0.24 Cal
NIH/NCI (PI: Vincent) $2,626,230
Gvl mHA Specific T Cell Responses Prevent AML Relapse Following Allogeneic Stem Cell Transplantation
We will use molecular subtypes to direct treatment in the neoadjuvant setting. To our knowledge, this is the first precision oncology trial in the neoadjuvant setting.

7. R01-CA244361 07/01/2020-06/30/2025 0.24 Calendar
NIH/NCI (PI: Thaxton) $1,250,247
Targeting Chronic ER Stress in T Cells to Improve Cancer Immunotherapy
To Identify radical new chronic ER stress targets that undermine the widespread success of immunotherapy in sarcoma patients and establish a new paradigm that informs drug development for all solid tumor cancer patients.

8. R01-CA241810 08/01/2020 – 04/30/2025 0.30 Cal
NIH (PI: Kim) $2,830,690
Chemotherapy and the Bladder Cancer Immune Microenvironment
The goal of this proposal is to assess how two different chemotherapy regimens with therapeutic equipoise, MVAC and GC, affect the immune microenvironment of bladder cancer as well as assess their ability to potentiate the effects of immune checkpoint blockade.

9. R01-CA248359 04/01/2020-03/31/2025 0.24 Calendar
NIH/NCI (PI: Thaxton) $1,026,742
Exploitation of ER Stress Induced Immune Dysfunction to Improve Immunotherapy
T cells are a part of the immune system able to control tumor growth, but are ineffective in most cancer patients. Our lab has made the groundbreaking discovery that T cells entering tumors experience an extreme response to stress that rapidly shuts down their ability to generate energy and make and process proteins. Our proposal will formally prove that targeting the stress response holds the power to revive essential T cell functions and invigorate tumor control to allow multiple types of immunotherapies to work better for cancer patients.

10. R01-CA229409 06/01/2019-05/31/2024 0.60 Cal
NIH/NCI (PI: Carey) $2,822,395
Optimizing HER2-Targeting Using RNA- and DNA-Based Predictive Algorithms
We propose to collectively integrate and analyze the clinical, gene expression, gene aberration, response to therapy, and outcomes data from more than 1500 women participating in multiple randomized neoadjuvant clinical trials of HER2-targeted therapy.

11. W81XWH2110693 09/01/2021 – 08/31/2024 0.32 Cal
Department of Defense (PI: Bryant) $657,868
Targeting KRAS-dysregulated metabolism for novel therapeutic approaches
Our application of metabolism-focused CRISPR libraries in orthotopic, syngeneic Kras-driven pancreatic mouse models will integrate tumor-TME interactions for more accurate modeling of the biological setting and therapeutic response of the patient cancer.

12. U24-AR076730 09/26/2019 – 05/31/2024 1.80 Cal
NIH/NIASD (PI: Lavange) $52,197,688
Back Pain Consortium (BACPAC) Research Program Data Integration, Algorithm Development and Operations Management Center
The goal of this proposal is to establish the Data Integration, Algorithm Development, and Operations Management Center (DAC) for the BACPAC Network at the University of North Carolina at Chapel Hill (UNC) Collaborative Studies Coordinating Center (CSCC) within the Gillings School of Global Public Health's Department of Biostatistics.

13. R01-CA230786 04/01/2019-03/31/2024 0.24 Cal
NIH-NCI (PI: Pylayeva-Gupta) $2,225,794
Function of IL35+ B cells in pancreatic cancer
Our major research goal is to understand immunosuppressive mechanisms that promote PDAC, which is the third leading cause of cancer death in U.S.

14. P50-CA058223-26 08/05/1997-08/31/2023 0.90 Cal
NIH-NCI (PI: Perou) $11,940,856
SPORE in Breast Cancer (Core B: Genomics, Biostatistics, and Bioinformatics)
The main goal of the Biostatistics and Bioinformatics Core (Core B) is to provide a complete and well-integrated Core for the analysis of complex multi-analyte data sets coming from our translational research of breast tumor specimens.

15. P50-CA058223-26 08/05/1997-08/31/2023 0.56 Cal
NIH-NCI (PI: Perou) $11,940,856
SPORE in Breast Cancer (Project 4)
The main goal is to identify the adaptive response to kinase inhibition in TNBC.

16. R21-CA246550 04/01/2020-03/31/2023 0.24 Calendar
NIH/NCI (PI: Dayton) $394,172
Parametric optimization of ultrasound-mediated immuno-modulation for pancreatic cancer therapy
Our research will use clinically meaningful pancreatic cancer murine models to provide defined ultrasound protocols for optimal modulation of PDA tumor microenvironment and test strategies that may enhance the impact of T cell-revitalizing therapies. This effort will inform the optimal design of ultrasound immunotherapy strategies against pancreatic cancer.

Completed

17. SAB180006 11/19/2018-11/18/2022 0.9 Cal
Susan G. Komen for the Cure (PI: Carey) $400,000
Optimizing HER2-Targeting Using RNA and DNA-Based Predictive Algorithms
We will examine the role of tumor and microenvironmental factors in determining response to HER2-targeting, relationship of pathologic complete response to outcome, and the biology of residual disease after dual or single HER2-targeting in HER2-positive breast cancer.

18. R01-CA199064 09/01/2016-07/31/2021 0.90 Cal
NIH-NCI (PI: Yeh) $3,587,399
Tumor Subtypes and Therapy Response in Pancreatic Cancer (Supplement RNASEQ)
This proposal was and continues to be based on our findings of two tumor-specific RNA subtypes in pancreatic adenocarcinoma patients and showed that these subtypes were consistently prognostic across independent datasets (ICGC and TCGA).

19. R01-CA199064 09/01/2016-07/30/2021 1.20 Cal
NIH-NCI (PI: Yeh) $3,587,399
Tumor Subtypes and Therapy Response in Pancreatic Cancer
This proposal was and continues to be based on our findings of two tumor-specific RNA subtypes in pancreatic adenocarcinoma patients and showed that these subtypes were consistently prognostic across independent datasets (ICGC and TCGA).

20. P01-CA203657 06/22/2016-05/31/2021 1.19 Cal
NIH/NCI (PI: Der) $7,857,810
Defining RAS Isoform- and Mutation-Specific Roles in Oncogenesis
The goal of this P01 program project is to complete a comprehensive study utilizing structural, biochemical and biological analyses using cell and mouse models of cancer to establish RAS isoform and RAS mutation specific functions in cancer.

21. P30-CA016086 12/1/15-11/30/20 3.60 Calendar
NIH/NCI (PI: Sharpless) $288,650.00
Cancer Center Core Support Grant: - Biostatistics
The principal objective of the Biostatistics SR Facility is to provide the highest level possible of quality statistical consultation services to UNC's Lineberger Comprehensive Cancer Center (LCCC) members.

22. V foundation Bae-Jump 11/01/17-10/31/20 0.6 Calendar
V foundation (PI: Bae-Jump) $600,000.00
Metabolic and Molecular Biomarkers of Metformin Response in Obesity-driven Endometrial Cancer
This study will test the hypothesis that higher rates of obesity and diabetes leads to disparate (higher) mortality in African American versus Caucasian endometrial cancer patients due to underlying biology using preclinical and patient samples.

23. U54-CA198999 8/1/15-7/31/20 0.6 Calendar
NIH/NCI (PI: Huang) $313,801.00
Nano Approaches to Modulate Host Cell Response for Cancer Therapy: Administrative Core
Administrative Core will coordinate all joint activities including advisory visits, speakers, reporting and other necessary task associated with the program

24. V foundation Major 12/01/2014-11/30/2018 1.2 Calendar
V foundation (PI: Major) $498,761
Team Science Approach for Defining the Activation State and Dynamic Reprogramming of the Kinome in Aerodigestive Cancer
Goals and Aims: We believe the proposed work, which couples cutting edge new technologies with ongoing clinical trials of targeted therapies, will significantly impact treatment and outcome for aerodigestive cancer patients. Aim 1 will determine the activation state of the

kinome in 50 primary human aerodigestive tumors; Aim 2 will correlate baseline kinome activity and kinome remodeling with innate and adaptive resistance to targeted therapies; Aim 3 will define the impact of targeted therapy on kinome activity using patient-matched pre- and post-treatment biopsies.

25. P50-CA058223 8/5/97-8/31/18 0.6 Calendar
NIH/NCI (PI: Earp) $187,370.00
SPORE in Breast Cancer - Project 2: Investigating the Function of the Immune Cell Infiltrate in the Biology of Claudin-low and Basal-like Breast Cancer
Project 2 will focus on generating CAR T cells from the PD-1 reporter mice to evaluate the activity of CAR T cells that express(ed) PD-1 in the tumor microenvironment.

26. P50-CA058223 8/5/97-8/31/18 0.6 Calendar
NIH/NCI (PI: Earp) $170,975.00
SPORE in Breast Cancer - Project 5: Defining Kinome Activity for Novel Therapies in Triple Negative Breast Cancer
Project 5 will focus on refining the mechanisms of enhancer induction by trametinib and further explore the functional regulation of the 186 genes whose common regulation by BRD4 and p300/CBP appear to be critical for adaptive bypass mechanisms in response to targeted kinase inhibitors

27. P50-CA058223 8/5/97-8/31/18 1.2 Calendar
NIH/NCI (PI: Earp) $146,476.00
SPORE in Breast Cancer - Core B: Genomics, Biostatistics, and Bioinformatics
The Genomics, Biostatistics, and Bioinformatics Core will provide valuable services for all projects. This Core will continue to explore new methods for the analysis of complex multi-analyte data sets, with an emphasis on data integration.

28. AACR Yeh 9/1/2016-8/31/18 0.6 Calendar
AACR (PI: Yeh) $228,156
Targeting macrophages to improve chemotherapy in metastatic pancreas cancer
The goal of this project is to provide the following support to the leading organization:1. Direct and analyze the bioinformatics data and samples of single cell core biopsies; 2. Classify tumor subtypes into single sample classifiers as per Dr. Yeh's Nature Genetics paper 3. Analyze pre- and post-therapy transcriptome changes in the tumor and stroma subtypes.

29. P50-CA058223 9/1/12-8/31/17 1.2 Calendar
NIH/NCI (PI: Earp) $142,502.00
SPORE in Breast Cancer - Core B: Genomics and Data Analysis
The Genomics and Data Analysis Core brings together the needed expertise and tools for analysis of multiple data types so that advances in breast cancer treatment can be made.

30. 2015YIA LEE 8/1/15-7/31/16 0.12 Calendar
Amer Society of Clinical Oncology (PI: Lee) $47,500.00
Combination CDK4/6 Inhibitor and MEK Inhibitor in KRAS Mutant Metastatic Colorectal Cancer
We propose to determine the efficacy of combination CDK4/6 and MEK inhibitors in patient-derived xenografts (PDXs) of RAS-mutant CRCs

31. P30-CA016086 12/1/10-11/30/15 3.60 Calendar
NIH/NCI (PI: Sharpless) $288,650.00

Cancer Center Core Support Grant- Biostatistics Shared Resource
The principal objective of the Biostatistics Shared Resource (BIOS SR) is to provide high quality statistical consultation services to UNC's Lineberger Comprehensive Cancer Center members. The BIOS SR provides a full collaborative scientific resource focused on providing Cancer Center members support for the design, conduct, analyses, and generation of manuscripts for their research.

SERVICE

Service within UNC-Chapel Hill

2023	UNC Biostatistics Faculty Retreat Planning Committee
2022–2023	Statistical Genomics Faculty Search Committee
2019-	Data Science Committee, Department of Biostatistics
2017-2019	Faculty Council, Gillings School of Global Public Health Representative
2017-	Genomics Joint Group Meeting (organizer), Department of Biostatistics
2016-2017	Statistical Genetics Faculty Search Committee, Department of Biostatistics
2016-	Masters Examinations Committee, Department of Biostatistics
2015-	Protocol Review Committee, Lineberger Comprehensive Cancer Center
2015-	Doctoral Examinations Applications Committee, Department of Biostatistics

Major External Service

2023-	Annals of Applied Statistics, Associate Editor
2023-	V Foundation for Cancer Research, Scientific Advisory Board member (invited)
2022-	NIH Cellular Immunotherapy of Cancer (CIC) Study Section, Statistical Reviewer
2020-	V Foundation for Cancer Research Grant Review Panel, Statistical Reviewer
2017-	Translational Breast Cancer Research Consortium, Statistical Working Group (invited)

Major Administrative Responsibilities

2022–	Associate Director: Biostatistics Core, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill
2022–2023	Chair, Statistical Genomics Faculty Search Committee
2021–	Chair, Doctoral Examinations Applications Committee, Department of Biostatistics, University of North Carolina at Chapel Hill
2018–	Associate Director: Cancer Genomics Training Grant, Department of Biostatistics, University of North Carolina at Chapel Hill

Ad Hoc Reviewer, Journals:

Journal of the American Statistical Association
Biometrika
Annals of Applied Statistics
JAMA Oncology
Clinical Cancer Research
Nature Breast Cancer
Journal of Clinical Medicine

PlOS Computational Biology
Genome Biology
BMC Bioinformatics
PLOS ONE
Genetics in Medicine