Jin Kweon, MSc

437-240-1343 | Montreal, Canada | yjkweon24@berkeley.edu | https://www.linkedin.com/in/jin-kweon/

SUMMARY

Analytical and results-driven data scientist with a strong foundation in data analysis and computational biology gained through a Master's in Experimental Medicine at McGill University and hands-on research experience. Skilled in developing scalable algorithms and statistical models to uncover biological insights. Seeking opportunities in data analytics, computational biology, or bioinformatics to apply expertise in Python, R, and machine learning to solve complex problems.

EDUCATION

McGill University   Montreal, QC
MSc, Experimental Medicine – GPA:4.00   Jan 2023 – Dec 2024
Thesis: "DTractor: Advancing cell type deconvolution of spatial transcriptomics with deep neural network, transfer learning and matrix factorization"

McGill University (Military service: 2019 – 2021)   Montreal, QC
PhD Candidate, Quantitative Life Science Withdrawn   Aug 2018 – Dec 2021

University of California, Berkeley   Berkeley, CA
Bachelor of Arts, Statistics (Focus: Mathematics / Economics) – Major GPA:4.00   Aug 2016 – May 2018
- **Relevant Coursework:** Applied Machine Learning, Structure and Interpretation of Computer Programs, Multivariate Statistics, Linear Algebra, Linear Modeling, Statistical Bioinformatics, Genomics for drug discovery and development

WORK & LEADERSHIP EXPERIENCE

Copoly.ai and Genentech   Toronto, ON
Computational Biologist   May 2025 – Present
- Developed novel statistical framework detecting disease-associated co-abundance at granular, single-cell resolution through differential correlation between cellular neighborhoods, advancing beyond CNA and Milo methods that test differential abundance within individual neighborhoods.
- Designed systematic interaction testing across cellular neighborhoods with genetics and spatial integration, revealing coordinated disease responses in AMD, IBD, AD, IPF, and SLE, including enterocyte-immune co-abundance signatures in Crohn's disease.
- Implemented computationally optimized statistical inference pipeline using FastQTL beta approximation and calibrated permutation testing with superior FDR control, enabling robust hypothesis testing across millions of neighborhood pairs.
- Pioneered compartment-based normalization with log-transformed abundance modeling to eliminate compositional bias, ensuring biologically meaningful comparisons across samples with heterogeneous cellular compositions.

Montreal General Hospital   Montreal, QC
Data Analyst   Feb 2025 – Present
- Pioneered ESDE 2025 award-winning methodology integrating multiple clinical databases (REDCap, TSQIC, ACCESS) establishing preoperative emotional distress as predictor of prolonged hospitalization and increased readmissions in esophageal cancer patient.
- Revealed in multivariable models that preoperative symptom burden predicted quality-of-life improvement one year post-esophagectomy, with smoking showing a possible trend, while tumor characteristics, surgical approach, and neoadjuvant treatment showed no association; female sex independently predicted reduced improvement.
- Led multi-year analysis (228 patients) demonstrating textbook outcome achievement as both a surgical quality benchmark and prognostic marker, correlating with significantly improved disease-free survival and favorable overall survival in esophageal cancer patients.
- Developed web-based scoring tools for FACT assessments to serve the broader surgical oncology community, and benchmarked 7 advanced imputation methods—including deep learning VAE/DAE and Bayesian PCA—to optimize missing data handling in oncology patient outcomes research.

Douglas Research Center   Montreal, QC
Computational Biologist   Mar 2025 – May 2025

- Boosted predictive accuracy from 55% to 95% by developing upgraded Python algorithms for meffil and ewaff R packages to detect differentially methylated positions by better leveraging non-linear relationships in the model between CpG methylation sites and phenotype variables.
- Engineered novel spline model methodology improving upon Enmix combp and bumphunter packages, successfully identifying previously undetectable sex-specific methylation regions and site changes across tissues and age.
- Created comprehensive model framework to analyze both sex-dependent and sex-independent methylation effects while ensuring robust evaluation of tissue-specific or tissue-consistent proliferation patterns over age.

Research on developing cell type deconvolution algorithms   Montreal, QC
Computational Biologist (spatial transcriptomics project)   Jan 2023 – Dec 2024
- Implemented the *DTractor* software that demonstrated higher accuracy, resolution, robustness, efficacy, and scalability on two simulated datasets and three real datasets, outperforming eight other popular and commonly used deconvolution methods.
- Developed a method that uses deep neural network training on sc/snRNA-seq and spatial transcriptomics (ST) data, incorporating transfer learning from scRNA-seq reference data to ST data, and conducting iterative matrix factorizations.
- Decomposed ST at cell-type resolutions to gain insights into spatial organization in pancreatic adenocarcinoma, uncovering tumor microenvironment interactions and offering rare capability to estimate cell counts per location.

Research on developing AI drug design-related algorithms   Montreal, QC
Research Analyst (single cell project)   Sep 2021 – Nov 2023
- Introduced *scBeacon*, an innovative framework built upon a VQ-VAE framework, deep contrastive siamese network, and a greedy iterative strategy, to effectively pinpoints differential genes and identify the same cell population across different biological conditions forming cluster pairs.
- Advanced biomarker discovery by liberating differential gene analysis from manual annotation constraints, enabling *scBeacon* to identify critical cluster pairs and potential therapeutic targets in COVID-19 datasets.
- Demonstrated therapeutic response by mapping cytarabine-induced cellular adaptations using optimal embedding vector quantization and probabilistic models, revealing elusive drug resistance pathways in treatment populations.

Adecco and Apple   Singapore
Apple Map Data Analyst   Jan 2022 – Dec 2022
- Optimized Apple Maps search accuracy in Korea through over 100 systematic evaluations of pedestrian routes, driving directions, search results, and auto-complete functions across iOS and macOS platforms.
- Resolved over 200 critical map bugs by implementing enhanced dataset analysis techniques, directly improving user navigation experience and reducing test-reported errors by 20%.
- Led development of comprehensive quality assurance protocols that measurably increased map data and query accuracy metrics across four key dimensions: location relevance, name verification, address precision, and pin placement.

Onbi, a smartphone application development company   Seoul, South Korea
Co-founder & CFO   Nov 2012 – June 2017
- Successfully led the "School Mom" project, an app designed to provide news and services to parents, resulting in our application being adopted by nearly 12,000 schools and attracting 50,000 users in Korea by conducting extensive surveys and researching trends and needs over several years.
- Organized an application that earned 2nd place in a South Korean National Founding Contest, with "School Mom" being recognized as influential in many schools nationwide.
- Managed diplomatic negotiations with major investors, securing a $100,000 contract through effective cold-calling.
- Achieved a net profit of over $10,000 within six months through meticulous budget management and consulting.

SKILLS
- Proficient in R [seurat], Python [pytorch, tensorflow, scanpy]. Basic knowledge of Java, SQL, CSS, HTML, MATLAB, Linux, C++, HPC/cloud computing, Git

PUBLICATIONS & ACKNOWLEDGEMENTS
- **Yong Jin Kweon**, Yousif Salman, Shayan Dhillon, Mehrnoush Dehghani, Emad A. Mohammed, R. Trafford Crump (2025). Comparing Missing Data Imputation Methods for Patient-Reported Outcomes in Esophageal Cancer Research. https://www.medrxiv.org/content/10.1101/2025.09.10.25335531v1

- Samir S. Amin, **Jin Kweon**, Joshua S. Cheruvathur, Mehrnoush Dehghani, Sara Najmeh, Carmen Mueller, Jonathan Cools-Lartigue, Lorenzo Ferri, R. Trafford Crump (2025). Preoperative Quality-of-Life and Recovery Following

Esophagectomy: Identifying Modifiable Factors Associated with Improvement. https://www.medrxiv.org/content/10.1101/2025.07.30.25332378v1

- **Yong Jin Kweon**, Chenyu Liu, Gregory Fonseca, Jun Ding (2025). DTractor enhances cell type deconvolution in spatial transcriptomics by integrating deep neural networks, transfer learning, and matrix factorization. https://www.biorxiv.org/content/10.1101/2025.04.12.648541v1
  GitHub: https://github.com/mcgilldinglab/DTractor

- Chenyu Liu, **Yong Jin Kweon**, Jun Ding (2023). scBeacon: single-cell biomarker extraction via identifying paired cell clusters across biological conditions with contrastive siamese networks. https://arxiv.org/abs/2311.02594
  GitHub: https://github.com/mcgilldinglab/scBeacon

- Yuying Lu, Camille Morencé, **Yong Jin Kweon**, Archer Yi Yang, Jun Ding (2025). SC2CD Co-optimizes Spatial Clustering and Cell Decomposition through Bidirectional Information Flow in Single-Cell Spatial Transcriptomics.

- Haber, J.R. (2020). Sorting Schools: A Computational Analysis of Charter School Identities and Stratification. Sociology of Education. https://doi.org/10.1177/0038040720953218

PRESENTATIONS

- **Jin Kweon**, Trafford Crump. (2025). Comparing Missing Data Imputation Methods for Patient-Reported Outcomes in Esophageal Cancer Research. Digital Health Symposium: From Research Infrastructure to Clinical Care. Montreal, Canada. Nov 4, 2025.

- Benjamin Tordjman, **Jin Kweon**, Nataliya Burban El Dirani, Mehrnoush Dehghani, Mathieu Rousseau, Carmen Mueller, Sara Najmeh, Jonathan Spicer, Jonathan Cools-Lartigue, Lorenzo Ferri, Trafford Crump. (2025). Evaluation of textbook outcomes and long-term prognosis following esophagectomy: a retrospective cohort study. International Society for Disease of the Esophagus (ISDE) World Congress. Brisbane, Australia. Sep. 18-20, 2025.

- Nisha Suarez, **Jin Kweon**, Mehrnoush Dehghani, Lorenzo Ferri, Sara Najmeh, Carmen Mueller, Jonathan Cools-Latrigue, Jonathan Spicer, Trafford Crump. (2025). Testing the effect of changes in emotional distress on post-esophagectomy outcomes. European Society for Diseases of the Esophagus (ESDE) conference. Amsterdam, Netherlands, May. 12, 2025. DOI: 10.17605/OSF.IO/UWBN2

- Samir Amin, Joshua Cheruvathur, **Jin Kweon**, Mehrnoush Dehghani, Sara Najmeh, Carmen Mueller, Jonathan Cools-Lartigue, Lorenzo Ferri, R. Trafford Crump. (2025). Rethinking Recovery Predictors in Esophageal Cancer Surgery. European Society for Diseases of the Esophagus (ESDE) conference. Amsterdam, Netherlands, May. 12, 2025. DOI: 10.17605/OSF.IO/QNHSE.

- Albouy, P., **Kweon, Y.J.**, Whittaker, H., Baillet, S., Zatorre, R.J. (2018). Enhancing learning-related plasticity with information-based neuromodulation. Auditory learning and plasticity symposium. Montreal, Canada, Oct. 23, 2018.