Dr. Yajing Zhu

Contact Information	Personalized Healthcare (PHC) Analytics, Roche Products Ltd, 6 Falcon Way, Shire Park, Welwyn Garden City, AL7 1TW, U.K. yajing.zhu@roche.com +44 (0)786 7134923

Summary

Passionate in leveraging big data to advance personalised healthcare and precision medicine. Statistician, entrepreneur, debater and consultant.

Positions

Senior data scientist, Neuroscience   Jan. 2021 -
Personalized Healthcare (PHC) Analytics, Roche

Visiting researcher   Jan. 2020 -
MRC Biostatistics Unit, University of Cambridge

Co-founder   Mar. 2020 -
Dx Analytics (currently in Accelerate Cambridge Ventures)

Education

London School of Economics and Political Science, London, U.K.

Ph.D., Statistics, passed with no corrections,   Nov. 2018
- Developed a multilevel (mixed-effects) model to relate clusters of individuals to multiple event histories and various health outcomes.
- Applied the methodology to a cohort study (N=19K,T=60 years). These analyses improved our understanding of the life course impact of social and health inequalities.
- Programming: R, STATA.

MSc., Statistics (Research), Distinction,   Oct. 2014
Dissertation: *The valuation of American option contracts.*
- Modelled the CBOE volatility index (VIX) using a jump diffusion (stochastic) process that improves the prediction of spot VIX.
- Programming: WinBUGS, R.

University of Liverpool, Liverpool, U.K
Xi'an Jiaotong-Liverpool University, Suzhou, China.

BSc.(Hons), Financial Mathematics, First Class,   Aug. 2013
Dissertation: *The valuation of American option contracts.*
- Compared three numerical schemes for pricing American put options.
- Wrote an efficient Matlab scripts for quick implementation with user-defined scenarios (simulation studies and empirical analyses).
- Programming: MATLAB.

Work experience

Co-founder of D-Art (predecessor of Dx Analytics)   Apr. 2019 -
Precision health check for artworks.
Finalist in the 3rd China (Shenzhen) Innovation& Entrepreneurship International Competition.

Research Associate, Precision medicine   Nov. 2018 - Dec. 2019
MRC Biostatistics Unit,University of Cambridge.
Multimorbidity research (electronic health records, N=400K, T=20 years).

Joint project with Addenbrooke's Hospital   Jul. 2019 -

Developed dynamic risk prediction models for ANCA-associated vasculitides with competing outcomes (relapse and infection).

Joint project with Papworth hospital   Mar. 2019 -

Developed a dynamic and personalised early warning tool using electronic health records of cardic surgical patients.

Improved sensitivity by 50% and tripled the precision of the current system.

Built a risk calculator (shiny app, R) with a doctor-friendly interface.

Worked with doctors and hospital researchers.

Joint project with DHSC of the UK government   Mar. 2016 -

Implemented a new method to compute the variance of margin estimates.

Built a Stata program (in mata) for efficient computation with user-defined constraints.

Worked with policy makers, PSNC and academics.

Visiting Researcher, Department of Geography   Sept. - Dec. 2014

Developed a statistical model to empirically test the migration framework in Burkina Faso.

Worked with geographers from University of Namur & University of Liege, Belgium.

REFEREED JOURNAL PUBLICATIONS

1. McClure*, M. E., Zhu*, Y., Smith, R., Gopaluni, S., Tieu, J., Pope, T., & Jones, R. B. (2020). Long-term maintenance rituximab for ANCA-associated vasculitis: relapse and infection prediction models. *Rheumatology*, 1-11.

2. Bendayan*, R., Zhu*, Y., Federman, A.D, Dobson, R. (2021), Multimorbidity patterns and memory trajectories in older adults: Evidence from the English Longitudinal Study of Ageing. *Journal of Gerontology: Medical Sciences*, In Press.

3. Zhu*, Y., Gleissl, T., Sanders, K., Squassante, L., Murtagh, L., Ahlers, S., Jan, A. and Smith, J. (2020). Clusters of Mental Health Comorbidities and Concomitant Medications of Participants in Three Randomized Controlled Trials (VANILLA, V1aduct, aV1ation) With Autism Spectrum Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, S162.

4. Zhu*, Y., Chiu, Y. D., Villar, S. S., Brand, J. W., Patteril, M. V., Morrice, D. J. & Mackay, J. H. (2020). Dynamic individual vital sign trajectory early warning score (DyniEWS) versus snapshot national early warning score (NEWS) for predicting postoperative deterioration. *Resuscitation*, 157, 176-184.

5. Zhu*, Y., Edwards, D., Mant, J., Payne, R. A., and Kiddle, S. (2020). Characteristics, service use and mortality of clusters of multimorbid patients in England: a population-based study. *BMC Medicine*, 18 (78).

6. Zhu*, Y., Steele, F., and Moustaki, I. (2020). A multilevel SEM for the interrelationships between multiple latent dimensions of childhood socioeconomic circumstances, partnership transitions and midlife health. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 183 (3), 1029-1050.

7. De Longueville*, F., Zhu*, Y., Henry, S. (2018) Direct and indirect impacts of environmental factors on migration in Burkina Faso: Application of Structural Equation Modeling. *Population and Environment*, 40(4), 456-479.

8. Zhu*, Y., Steele, F., Moustaki, I., (2017). A General 3-Step Maximum Likelihood Approach to Estimate the Effects of Multiple Latent Categorical Variables on a Distal Outcome. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(5), 643-656.

9. Zhu*, Y. and Cheng, J., (2013). Using hidden Markov model to detect macroeconomic risk level. *Review of Integrative Business and Economics Research*, 2(1), 238-249.

PAPERS IN PREPARATION	1. Can Multidimensional Poverty Index Substitute Income to Target the Poor? A Study Based on the Poor Household Database of a Nationally Designated Poor County in China.
	2. Organizational Antecedents of Evil: The Effect of Supervision on Workplace Abuse.
	3. Trajectory Modelling in Preclinical Alzheimer's Disease: Insight from ADNI, Biofinder-1 and AIBL.

SELECTED PRESENTATIONS

- **(Invited workshop)** Workshop on Statistical methods for Cluster Analysis applied to multimorbidity.
  Statistical Methods for Multi-Morbidity Community, Virtual meeting.   Nov. 2020
- **(Invited talk)** A novel method to identify clusters of multimorbid patients in linked CPRD primary and secondary care EHR.
  Medicines and Healthcare products Regulatory Agency, London   Jul. 2020
- **(Invited talk)** Relating clusters of patients to distal outcomes: A misclassification-correction approach.
  Computational and Methodological Statistics, London   Dec. 2019
- **(Invited workshop)** Understanding the life course impact of childhood socioeconomic disadvantages on midlife health: a multilevel SEM approach.
  Institute of Child Health, University College London   Jan. 2019
- Characteristics, service use and mortality of clusters of multimorbid patients in England: a population-based study.
  The Lancet Public Health Science 2019, London   Nov. 2019
- Identification of clusters and outcomes of multimorbid patients in UK general practice.
  12th European Public Health conference, Marseille   Nov. 2019
- A 3-step multilevel SEM
  Statistical Analysis of Multi-outcome Data, Manchester   Jun. 2019
- Extending the 3-step approach to a multilevel structural equation model.
  2018 International Meeting of Psychometric Society, New York   Jul. 2018
- Application of Structural Equation Modeling for analysing impacts of environmental factors on migration in Burkina Faso.
  2017 International Population Conference, Cape Town   Nov. 2017
- An extension of the general 3-step ML approach to random effects EHA with multiple latent categorical predictors.
  Royal Statistical Society Annual Conference, Glasgow   Sept. 2017
- A general 3-step method to estimate the effects of multiple latent categorical predictors on a distal outcome.
  Royal Statistical Society Annual Conference, Manchester   Sept. 2016

GRANTS/AWARDS **2020**
- Two-year research grant on "Dynamic Prediction Methodology: with applications in Neuroscience and Oncology" (£210,000). Joint fund by Roche/Genentech and University of Cambridge, MRC Biostatistics Unit.

2015 - 2019
- Associate Fellow of The Higher Education Academy, UK.
- LSE Travel grants (Aug. 2018, £1,000; Jul. 2017, £1000)
- LSE Best Class Teacher Award, London School of Economics.
- LSE Statistics PhD Scholarship (£120,000), London School of Economics.

2009 - 2013
- National Scholarship (Ministry of Education of the P.R China).
- National Endeavor Scholarship (Ministry of Education of the P.R China).
- Provincial Outstanding Student in Jiangsu Province, P.R China.

- University first-class progression scholarship (3 years in a row).
- Departmental Best Overall Performance (Dept. of Mathematical Science).
- Honourable Mention in Mathematical Contest in Modelling (The Consortium for Mathematics and its Applications, U.S.)
- 'Best Debater' and 'Winning Team' in the Debating Society.

TEACHING EXPERIENCE	Class teaching at London School of Economics, U.K.

Dept. of Statistics
- ST411 Generalised linear modelling   2017 - 18
- ST405 Multivariate Methods   2017 - 18
- Elementary Statistical Theory   2015 - 17

Dept. of Economics
- LSE Summer School 2017: Essential Statistics for Economics and Econometrics
- LSE Summer School 2017: Further Statistics for Economics and Econometrics

SERVICES
- Departmental PhD representative.   2015 - 19
- Athena Swan EDI Taskforce: Student representative.   2016 - 19

SOFTWARE AND LANGUAGE SKILLS

Computer Programming: R, Stata, Matlab, WinBUGS, Mplus, LatentGOLD

Language: English (fluent), Mandarin (native), French (beginner).