林　汇亨　***Hui-Heng Lin***　**Ph.D.** in Systems and Computational Biomedicine
Linkedin: www.linkedin.com/in/res000/　**ORCID:** orcid.org/0000-0003-4060-7336
Github: https://github.com/devgithwaar?tab=repositories
Email: lin.hui.heng@connect.um.edu.mo　**TEL:** +86 17138862101

Expertise

Computational Analytics & Predictive Modelling | Systems & Computational Biomedicine | Interdisciplinary Researches | Network Sciences

Skills

Statistical analyses / Predictive modelling / Machine learning / Bioinformatic databases & toolkits / Complex social network analysis / Raw data processing and structuralizing / Reporting & data visualizations / Scientific English writing and editing

Technologies

Linux & Shell scripting / Python / R / SQL / Microsoft Excel Advance Functions / SPSS

Highlighted Research Projects (*More details on my websites*)

Project # 1 **Machine learning prediction** on BRCA1 gene variants' cancer pathogenic risks.

(1) Using programming codes and regular expression, I collected and *transformed unstructured 6k+* DNA sequences and other types of raw data into *structured* datasets. And further generated numeric representations for *text / character / string data*.

(2) *In-house* trained multiple machine learning models including *naïve bayes, support vector machine, logistic regression, random forest.* Benchmarked their predictive performances with external reference models (Polyphen and SIFT). *My best model outperformed reference models* ***upto 11 %*** in terms of area under curve.

(3) Performance of the random forest model ***increased from 80% to 85%*** (area under curve) after optimization. And it was used to *predict the cancer pathogenic risk of 10 clinical patients* (7 oncogenic BRCA1 gene variants with unknown/unsure pathogenicity), *successfully facilitated diagnoses of patients.*

(4) Data and analytic *results were plotted, visualized* and reported. The project was summarized to ***my first-authorship*** *publication* (see Lin HH, et al. Predicting ovarian/breast cancer pathogenic risks of BRCA1 gene variants of unknown significance. *MedRxiv* 2020. https://doi.org/10.1101/2020.06.04.20120055)

(5) The project also *won research grant* from Shaoguan Science and Technology Bureau (Grant ID: 200812114531428). *Hui-Heng Lin acts as the* ***Principle Investigator*** of this project.

(6) Techs used: Self-scripted R codes and libraries (Biostrings, Bioconductor, e1071, randomForest, protr, pROC, ROCR, ggplot2, etc)

Project # 2 **Complex network modelling & link prediction** for computational **drug repositioning** (Anti-cancer drugs were tried to be repurposed for neurodegenerative diseases such as Alzheimer's and Parkinson's disease)

(1) Used molecular descriptor algorithms (fingerprint and atom pair descriptor) to analyze the chemical ***structure similarity of 40,000 pairs of drugs***. Cosine-similarity index was used to measure the ***phenotypic similarity of 1024 pair*** neurodegenerative diseases. Both similarity score (Jaccard similarity coefficient & Cosine-similarity index) matrices were *visualized into heatmaps*.

(2) Through ***data normalization, standardization*** *and integration*, an *adjacency matrix* of the drug-disease bipartite graph was obtained. The matrix was *visualized to a heterogeneous bipartite drug-disease network* with 232 nodes, and 1007 edges. Network ***topological and modularity analyses*** *were conducted.*

(3) 8 types of link prediction ***(unsupervised learning)*** *algorithms*, including *clustering / community / modularity-based* and eigenvector-based link predictors were applied to the network, and *leave-out-one metric was developed to benchmark* the average performance of link predictors. Performances of predictors were sorted and *visualized into Boxplot (or violin plot)*.

(4) Rooted PageRank was identified as the best link predictor and it was used to ***predict 6367 anti-cancer drug—neurodegenerative disease*** *potential therapeutic association edges* in the network. Drug-disease pair edges of top-ranked link prediction scores were queried against literature database, and within 7 drug-disease pairs of top-hit, *3 anti-cancer—neurodegenerative diseases interaction were found to be supported by research publications*.

(5) Data and analytic *results were plotted, visualized* and reported. The project is being summarized to research article for academic publication.

(6) The proposal of this project also *won Ernst Mach grant / fellowship* (Reference ID: ICM-2018-10230) from EurAsia pacific network foundation/OeAD Austria.

(7) Techs used: Self-scripted python codes and modules (igraph, networkX, numpy, pandas, RDkit, matplotlib, linkprediction), R (igraph, SANDS, LinkPrediction, ChemmineR, dSimier, pheatmap, disgen2r, Bioconductor, ggplot2), gephi, etc.

Project # 3 **Statistical and Bayesian** analyses for survey data on clinical healthcare.

(1) Questionnaires were designed for surveying 110 clinicians regarding healthcare issues. ***Psychometric properties*** of the questionnaire including ***reliability (via Cronbach alpha reliability test)*** and ***validity (via Factor analysis. Kaiser-Meyer-Olkin and Bartlett's tests)*** were confirmed. Clinicians' answers were sorted and character data were transformed into numeric score data, followed by ***statistical analysis***.

(2) ***2860 times of Chi-squared*** tests were conducted to ***detect the potential association (likelihood)*** between hospitals' cost-control polices and healthcare issues. The tests' data *(P-value < 0.01)* were used to construct a ***Bayesian (probabilistic) network*** with 25 nodes and 43 edges. ***Multivariate analyses and Bayesian probability inferences*** were carried out to ***identify the cause-effect relations*** between network nodes (hospitals' cost-control polices and healthcare issues).

(3) Data and analytic *results were plotted, visualized* and reported. The project has been published as "Yan J, ***Lin HH****, et al., China's new policy for healthcare cost-control based on global budget. *BMC Healthcare Service Research*, 2019, 19:84. https://doi.org/10.1186/s12913-019-3921-8 (****Hui-Heng Lin*** *was the* ***co-first author*** of this research work and article, and *substantially contributed in project co-ordinations, design of research methodologies, data analysis and statistical modelling, English manuscript, supplementary materials, English sample questionnaire, drafting, revisions, and reviewing)*.

(4) This research work was *recognized by Journal of quality in healthcare & economics, Journal of health economics and outcomes research, Journal of health policy and outcomes research*, etc.

(5) Techs used: R (igraph, ggplot2, pheatmap, etc); gephi, SPSS, Microsoft Excel, etc.

Research & Working Experience

2019 ~ Present*:* ***Associate Investigator / Medical Scientist*** for computational medical sciences (Postdoctoral)

@ Yuebei people's hospital, *Shantou University Medical College affiliated* (at Shaoguan, China*)*

1. **Researches**: Independently research on projects on *machine learning clinical bioinformatics* and translational medicine. Writing and publishing research article in academic journals
2. **Grant acquisition**: Drafting research proposals for research grant applications. *Acquired research grants* of internal and external funder (Shaoguan Science and Technology Bureau. Grant ID: 200812114531428. Role: project leader and principal investigator)
3. **Awards**: *Awarded fellowships* from Shaoguan city and Guangdong province (postdoctoral)
4. **Supervisions:** *Cosupervised 3 master students'* clinical research projects thesis writing. Advised on designs of project workflows and methodologies. Projects include clinical meta-analysis, retrospective studies, and statistical/predictive modelling of clinical (patients, survivals, gene expressions, etc.) data analysis.
5. **Collaborations**: *Led project members in multiple grant applications*. *Initiated collaborative research* projects with different clinical and therapeutic departments in hospital. E.g., a network pharmacological analytic project on traditional Chinese medicine as the complementary treatment for cancers, and clinical screening and analysis of patients' genetic mutations and gene variants.
6. **Consultations**: Provided professional research advice to a research project about network pharmacology of traditional Chinese medicines, which was initiated by hospital's department of pharmacy.

2015 ~ 2019: **Graduate Research Fellow**

@ State key laboratory, Institute of Chinese Medical Sciences, University of Macau (at Macau, China)

1. **Researches**: Independently conducted multiple biomedical and *interdisciplinary* research projects using *bioinformatics* analytics, *machine learning predictive* modelling, social / complex *network analytics*, and *statistical* modelling approaches.
2. **Academic publishing:** *High productivity* in scientific publications. Successfully in publishing *4 first / co-first authorship* original research articles in Science-Citation Indexed academic journals and multiple research chapters in monographs (see my **ORCID:** https://orcid.org/0000-0003-4060-7336 **).**

3. **Fellowships & Grants: (1)** *Ernst Mach Fellowship & Grant* from Eurasia Pacific Network foundation and OeAD Austria (Reference ID: ICM-2018-10230). Academically visted to Austrian research institutions and conducted collaborative researches there during year 2018 ~ 2019. **(2)** *Hong Kong Polytechnic University Joint Supervision Scheme* during year 2017 ~ 2018 (Fellowship & grant ID: G-SB81)
4. *Completed 3 collaborative* research projects with internal and external collaborators and *co-supervised* master students' research projects and paper / thesis writing.
5. **English presentation award** in University of Macau Institute of Chinese Medical Science, academic forum (2016)
6. **Community service**: Selected and severed as the residential tutor and committee of college discipline of Lui Chi Woo College at University of Macau **(**Year 2016 ~ 2017. University community service**)**
7. **Teaching** (in full English)**:** Research methodologies. Section of Introduction to research in medical administration (2017 Fall semester, Class code CMED710).

2011 ~ 2015: **Research Assistant**

@ University of Tokyo (at Tokyo, Japan)

1. Conducted independent biomedical research projects and assisted research projects of other lab members
2. Lab managements including lab safety checking, preparations and managements of biochemical agents and materials, and maintenance of lab equipment, instrumental devices and biological/chemical stocks.
3. Biochemical skillset: Gene cloning & mutagenesis, Cell/Bacterial culture, DNA sequencing, Protein expression & purification by liquid chromatography, Surface plasmon resonance (Biacore), etc.

Education

2015 ~ 2019: **Ph.D.** (Excellence) in Systems & Computational Bio-medicine

- State key laboratory, Institute of Chinese Medical Science, *University of Macau (*At Macau, China)
- Thesis: Repositioning anti-cancer agents for neurodegenerative diseases via drug-disease network analysis
- Core curricula: Advanced statistical modelling / Graph theory & Complex networks / Social & complex network analytics / Systems biomedicine / Network poly-pharmacology / Research methodologies / Scientific writing & reporting in English

2012 ~ 2014: M.Sc. (Excellence) in Medical Sciences (Research expertise: Molecular oncology)

- Graduate school of medicine, *University of Tokyo* (At Tokyo, Japan)
- Core curricula: Oncology / Immunology / Medical & infectious microbiology / Biophysics & biochemistry / Genomics & Bioinformatics
- Thesis: in vitro reconstitution of oncogenic protein CagA-Csk kinase interaction (Grade: Excellent)

2007 ~ 2011: B.Sc. in Biotechnology (Top-ranked 10%)

- School of Biosciences, *South China University of Technology* (Project 985 National Key University, At Guangzhou, China)
- Core curricula: Computer programming & technologies / Linear algebra & probability statistics / Differential & calculus / Bioinformatics & computational biology / Chemical sciences / Microbiology / Molecular & cell biology / Genetic & fermentation engineering
- Thesis: Molecular modelling on 3D structure of Zirconium compound family

Language Skills

Native Cantonese & Chinese Mandarin \ **Fluent English:** International Business & academic research proficiency (***TOEFL-iBT score 99*** & College English Test-Band 6 certificate) \ **Japanese:** *Japanese Language Proficiency Test Level 1 certificate (N1)* \ Basic **German.**

Research Publications in Academic Journals

1. **Lin HH**, Xu H, Hu H, Ma Z, Zhou J, Liang Q. predicting ovarian/breast cancer pathogenic risks of BRCA1 gene variants of unknown significance. *MedRxiv* 2020.
2. **Lin HH.** Repositioning anti-cancer agents for neurodegenerative diseases based on drug-disease analysis. University of Macau Ph.D. graduation dissertation thesis 2019.
3. Wang Z, **Lin HH***, Linghu K, Huang RY, Li G, Zuo H, Chan G, Hu Y. Novel compound-target interactions prediction for the herbal formula Hua-Yu-Qiang-Shen-Tong-Bi-Fang, *Chemical and Pharmaceutical Bulletin.* 2019. (*Co-first authorship, **Science citation-Indexed**)
4. Yan J, **Lin HH***, Zhao D, Hu Y, Shao R. Deterioration of the health quality: a study on healthcare cost control policy and actions in China. *BMC health service research*. 2019, 19:84. DOI:10.1186s12913-019-3921-8 (*Co-first authorship, **Science citation-Indexed**).

5. **Lin HH**, Ouyang D, Hu Y. Intelligent Classifier: a Tool to Impel Drug Technology Transfer from Academia to Industry. *Journal of Pharmaceutical Innovation*. 2018:1-7. (**Science citation-Indexed**)
6. Liu K, **Lin HH**, Pi R, Mak S, Han Y, Hu Y. Research and development of anti-Alzheimer's disease drug: an update from the perspective of technology flows. *Expert Opinion On Therapeutic Patents.* 2018 (**Science citation-Indexed**).
7. **Lin HH**, Zhang LL, Yan R, Lu JJ, Hu Y. Network Analysis of Drug–target Interactions: A Study on FDA-approved New Molecular Entities Between 2000 to 2015. *Scientific reports*. 2017 Sep 25;7(1):12230.
8. Huang C, **Lin HH***, Wan JB, He C, Hu Y. Research and Development of Hepatitis B Drugs: An Analysis Based on Technology Flows Measured by Patent Citations. *PloS one*. 2016 Oct 11;11(10):e0164328. (*Co-first authorship, **Science citation-Indexed**)

Publications in Monographs / Book Chapters

1. Huang C, **Lin HH***, Wan J, He C, Hu Y. Research and Development of Hepatitis B Drugs: An Analysis Based on Technology Flows Measured by Patent Citations. *Innovation Value Chain of Chinese Pharmaceuticals*, United States-China Intellectual Property Institute, 2017, Chapter 8. (*Co-first authorship)
2. Tian S, **Lin HH***, Mak MSH, Han Y, Hu Y. Analysis of Technology Flows of Antidepressants Based on Patent Citation Network. *Innovation Value Chain of Chinese Pharmaceuticals*, United States-China Intellectual Property Institute, 2017, Chapter 9. (*Co-first authorship)
3. Li P, **Lin HH**, Hu Y. Pharmaceutical Innovation Network: Global Patterns and the Role of China. *Innovation Value Chain of Chinese Pharmaceuticals*, United States-China Intellectual Property Institute, 2017, Chapter 13.
4. **Lin HH,** P Li, Kong XJ, Cai J, Su SB, Hu Y. An analysis of evolutionary conservation based on genomic orthology and protein-protein interaction networks. *Biotechnology in Hong Kong III*. United States-China Intellectual Property Institute, 2017. Chapter 4.