Lucas V.H.H-TRAN, PhD

Biomarkers Data Scientist with 10 years of experience and a Bachelor's degree in Pharmacy, driving impactful insights in healthcare and pharmaceutical research.

Age: 35 ✻ *Nationality:* French✻ *Swiss work permit:* B

Residence: Davos Platz, Switzerland

E-mail: tranhungydhcm@gmail.com ✻ *Telephone:* +41 76 483 31 83

Links: github.com/vanhungtran ✻ Γ 0000-0002-5226-3553 ✻ *Website:* vanhungtran.github.io

Work Experience

Christine Kühne – Center for Allergy Research and Education *Nov 2022 – Present*
Biostatistician (Observational Study & Preclinical) *Davos, Switzerland*

- Managed and analyzed data for over 3,000 atopic dermatitis patients in the PRORAD study, ensuring adherence to ICH/GCP standards.
- Collaborated with biologists, clinicians, and engineers to deliver actionable insights, enabling more informed research directions and decisions.
- Served as Statistical Project Manager for Davos Biosciences, a non-profit Biotech SME and spin-off of CK-CARE; supervised students.

AXIA *Apr – Oct 2022*
Preclinical Oncology Biomarker Biostatistician (Consultant at SANOFI) *Paris, France*

- Applied statistical and machine learning methods for target identification, biomarker prediction, and pathology analysis in drug discovery.
- Reviewed statistical analysis plans and developed corresponding statistical programs for implementation.
- Developed tools to analyze multi-omics data from oncology clinical trials.
- Prepared Spotfire dashboards for reporting and visualizing deliverables.

ENDODIAG *Jun 2020 – Mar 2022*
Biomarker Data Scientist (Preclinical Phase) *Paris, France*

- *ENDODIAG develops novel diagnostic solutions for endometriosis to improve patient management, treatment, and fertility strategies.*
- Performed data management (obtaining, merging, cleaning, and quality control) in accordance with Good Clinical Practice (ICH/GCP).
- Handled missing data using multiple imputation and deep learning techniques.
- Analyzed microRNA, RNA-seq, and proteomics data for biomarker discovery using R/Bioconductor and Python.
- Built machine learning pipelines (feature engineering, selection, and optimization) to identify transcriptomic and proteomic signatures for endometriosis prediction.
- Utilized PyCaret, H2O, scikit-learn, TensorFlow, and PyTorch to model multi-level clinical trial data.
- Created intuitive graphics and visualizations for data analysis and reporting.

Bordeaux School of Public Health *Nov 2018 – Dec 2019*
Postdoctoral Biostatistician (Phase 1-2 Clinical Trials) *Bordeaux, France*

- *Contributed to the European project DALIA, investigating a dendritic cell vaccine as a safe and effective treatment for HIV patients.*
- Applied knowledge of clinical trial design, regulatory requirements, and causal/Bayesian inference.

- Developed machine learning and statistical tools to analyze longitudinal data from Phase 1-2 HIV vaccination trials.
- Compared clustering approaches (cytometree, flowMeans) for quantifying immune cell subgroups from flow cytometry data.
- Developed R packages and Shiny Apps for data visualization and analysis.

National Research Institute for Agriculture, Food and Environment *Oct 2015 – Oct 2018*
Research Assistant (EU Project Feed-a-gene) *Toulouse, France*

- *Developed novel statistical models based on linear mixed models to handle changes of (genetic) random effects over time.*
- Modeled longitudinal data using time series analysis, forecasting, and linear/non-linear mixed models.
- Estimated variance components for longitudinal data using random regression and structured antedependence (SAD) models with ASReml and BLUPf90.
- Evaluated the potential of genomic information in predicting phenotypes (GWAS, single step).

University Hospital of Lyon *Jan 2015 – Jul 2015*
Research Assistant *Lyon, France*

- Analyzed longitudinal Aspergillus antigen data in immunocompromised patients using imputation for missing data.
- Developed survival and prognostic models to create a predictive score for early detection of Invasive Aspergillosis.

Education

National Polytechnic Institute of Toulouse, France *Oct 2015 – Oct 2018*
Ph.D. in Statistical Methods for Quantitative Genetics
Thesis: New longitudinal genetic models for feed efficiency, theses.hal.science/tel-02789358v1

University Lyon 1, France *Sep 2014 – Aug 2015*
M.Sc. in Biostatistics, Biomathematics, and Bioinformatics
Thesis: Survey of Serum Aspergillus Antigen in Patients [...] Role in Predicting Invasive Aspergillosis, DOI:10.1182/blood.V126.23.3428.3428

Hochiminh University of Medicine and Pharmacy, Vietnam *Sep 2008 – Aug 2014*
Pharm.B.
Thesis: Ex-Vivo Percutaneous Absorption of Enrofloxacin: Comparison of LMOG Organogel vs. Pentravan® Cream, DOI: 10.1016/j.ijpharm.2015.12.018

Projects

Ebook	Large Language Models-FAQ (bookdown.org)
Post	Rpubs (https://rpubs.com/hung/)
Software develop	BiomartScope, aucmat, atcddd, RNAseqDataDownloader (github), MLFeatureSelection (shinyapps.io 1), FDA novel approval(shinyapps.io 2)
Multi-omics Data Analysis	CK-CARE, SANOFI,JNJ, UCB(Data analysis Report)
ML for Biomarker Discovery	ENDODIAG (endodiag.com)
Longitudinal Genetic Models	New genetic models for feed efficiency (PhD Thesis)
Flow Cytometry Data Analysis	NK Cell Receptor Repertoire Analysis (Frontiers in Immunology)
Statistics in Diagnostics	Survey of Serum Aspergillus Antigen (Blood Journal)
Kinetic Modeling	Ex-Vivo Percutaneous Absorption of Enrofloxacin (Int. Journal of Pharmaceutics)

Technical Skills

Statistical/ML Languages	R, Python, SAS
Other Programming	Julia, LaTeX, Shell, Fortran, Matlab, HTML/CSS/JS
Specialized Tools	ASReml, BLUPf90 family, SciLab
Cloud Platforms	AWS, Azure, GCP
Environment	Windows, Linux, macOS, Git, GitLab CI, Jupyter Notebook
Analytics Platforms	Tableau, Spotfire

Languages

English	Professional Working Proficiency
French	Professional Working Proficiency
German	Basic Proficiency

REFERENCES

Dr. Wenting Wang
Global Head of Biomarker Statistics
SANOFI
✉ Wenting.Wang2@sanofi.com

Dr. Hélène Gilbert (PhD supervisor)
Research director
INRAE Toulouse
France
✉ helene.gilbert@inrae.fr