Lucas V.H.Huynh TRAN, PhD

Age: 20+ ✻ *Marital status:* Single

Nationality: French ✻ *Swiss work permit:* B

Residence: Davos Platz, Switzerland

E-mail: tranhungydhcm@gmail.com ✻ *Telephone number:* +41-764-xx-xx-xx

Github: https://github.com/vanhungtran ✻ *ORCID:* https://orcid.org/0000-0002-5226-3553

Work experience

Christine Kühne – Center for Allergy Research and Education   *Nov 2022 - present*
Biostatistician/Bioinformatician (Observational Study & Preclinical Phase)   *Davos, Switzerland*

- Data management of 3000 atopic dermatitis patients - PRORAD study, according to good clinical practice (ICH/GCP).
- Support study design and develop and implement data analysis pipelines for multi- omics data.
- Collaborated with cross-functional teams of biologists, clinicians, and engineers to translate the findings into actionable insights and solutions.
- Statistical project manager for Davos Biosciences (a non-profit Biotech SME and spin-off of CK-CARE) and supervise students.

SANOFI   *April - Oct 2022*
Preclinical, 1-2-Phase Oncology Biomarker Biostatistician/Data scientist (Contractor) Paris, France

- Leveraging statistical and machine learning approaches in drug discovery and development includes target identification and validation, small-molecule design and optimization, prediction of biomarkers, and computational pathology.
- Review the statistical analysis plan and develop statistical programs
- Develop machine learning and statistical tools for analyzing omics data obtained from clinical trials on oncology.
- Preparing dashboards with Spotfire as deliverables.

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

- *ENDODIAG, a company based in Paris, France, is developing new endometriosis diagnostic solutions for healthcare professionals. These solutions are intended to contribute to a better endometriosis diagnosis, more personalized patient management, more efficient treatment options, and fertility strategy.*
- Obtain, merge, reshape, clean, ensure, and sustain data quality according to good clinical practice (ICH/GCP).
- Review the statistical analysis plan and develop statistical programs.
- Handle missing data with multiple imputation and deep learning.
- Use descriptive statistics, predictive analytics, machine learning, and other methods to learn about and derive insights from the patterns and relations within multi-omics data.
- Analyze microRNA (coding and non-coding), RNAseq, and proteomics expression analysis for biomarker discovery with R Bioconductor and BioPython.
- Perform machine learning (including deep learning) with PyCaretlow, AutoML (H2O), sklearn, TensorFlow, keras, Torch. . . to explore and model data obtained from clinical trials (survey data, genomics, transcriptomics, proteomics, pathology images, ...).

- Assess transcriptomic and proteomic signatures to predict endometriosis using different statistical and machine learning approaches (pipeline for feature engineering, feature selection, generate marker combinations, and optimize them to obtain the best signature with given biomarkers or/and according to predefined objectives).
- Create intuitive and compelling graphics to visualize and think about data, as well as for reporting.

Bordeaux School of Public Health
Postdoctoral Biostatistician (Phase 1-2 clinical trials)

Nov 2018 – Dec 2019
Bordeaux, France

- *This is part of European project DALIA(HIV Vaccination). The purpose of this study is to determine whether the administration of a dendritic cell vaccine is a safe and effective treatment for HIV-1 patients.*
- Knowledge of the design of clinical trials and regulatory requirements.
- Univariate and multivariate statistical analysis.
- Develop machine learning and statistical tools for analyzing longitudinal data obtained from HIV vaccination Phase 1-2 clinical trials.
- Compare different approaches: cytometree, flowMeans, Flowmerges, and Flowcyto for determining and quantifying different subgroups of immune cells.
- Statistical inference: Bayesian, causal inference.
- Develop a package, Shiny Apps, and a dashboard with R.
- Able to work with Monolix (software in pharmacometrics).

National Research Institute for Agriculture, Food and Environment
Research assistant (EU Project Feed-a-gene)

Oct 2015 – Oct 2018
Toulouse, France

- *This is part of the European Feed-a-gene. The purpose of this study is to develop new statistical models based on linear mixed models that are able to handle the changes of (genetic) random effects over time.*
- Modeling longitudinal data, time series analysis, and forecasting.
- Use linear and non-linear mixed models to handle missing data
- Analysis of variances, matrix decomposition, parameterization, and developing models to estimate variance components for longitudinal data on the framework of random regression and structured independence models (with ASReml, BLUPf90 (GWAS, single step)).
- Evaluate the potential of genomic information in predicting the phenotypes.
- Apply machine learning approaches for clustering and identifying different genetic profiles over time.
- Develop Structure Antedependence (SAD) models and implement them in ASREML.

University Hospital of Lyon
Research assistant

Jan 2015 – July 2015
Lyon, France

- Use an imputation to correct the missing data.
- Analysis of Aspergillus antigens (longitudinal data) in immunocompromised patients.
- Survival models and prognostic models.
- Propose a score for early prediction of the Invasive Aspergillosis

Education

National Polytechnic Institute of Toulouse, France   *Oct 2015 - Oct 2018*

Ph.D. in Pathology, Toxicology, Genetics, and Nutrition (Statistical Methods for the Quantitative Genetics)

Thesis: New longitudinal genetic models for feed efficiency, https://theses.hal.science/tel-02789358v1

University Lyon 1, France   *Sep 2014 - Aug 2015*

M.Sc. in Biostatistics, Biomathematics, Bioinformatics, and Health

Thesis: Survey of Serum Aspergillus Antigen in Patients Undergoing Chemotherapy for Acute Myeloid Leukemia or Allogeneic Hematopoietic Stem Cell Transplantation: Role in Predicting Invasive Aspergillosis and in Modifying the Therapeutic Strategy, DOI: 10.1182/blood.V126.23.3428.3428

Hochiminh University of Medicine and Pharmacy, Vietnam   *Sep 2008 - Aug 2014*

PharmB

Thesis (Internship at University Lyon 1, France): Absorption of Enrofloxacin: Comparison of LMOG Organogel vs. Pentravan® Cream EA 4169 Fonctions Physiologiques et Pathologiques de la Barrière Cutanée, 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
Multi-omics data analysis	CK-CARE, SANOFI, JNJ, UCB
Machine learning for Biomarker Discovery	ENDODIAG
Statistics in diagnostic medicine	Survey of Serum Aspergillus Antigen
Flow cytometry Data Analysis	Distinct Signatures in the Receptor Repertoire Discriminate CD56bright and CD56dim Natural Killer Cells
Longitudinal genetic models	New genetic longitudinal models for feed efficiency (joint genotypic and phenotypic outcome modeling)
Kinetic Modeling	Ex-Vivo percutaneous absorption of enrofloxacin: Comparison of LMOG organogel vs. pentravan cream

Technical skills

Statistical/ML Programming Languages/Tools	R, SAS, Python
Natural Language Processing	R, Python
Other Programming Languages/Tools	LaTeX, Shell scripts, ASReml, Julia, BLUPf90 family, SciLab, Fortran, Matlab, HTML/CSS/Javascript
high-performance computing cluster	AWS, Azure, GCP
Environnement	Window, Linux, Mac, Git, Gitlab CI, Jupyter notebook
Analytics Platform	Tableau, Spotfire

Languages

English	Professional Working Proficiency
French	Professional Working Proficiency
German	Basic