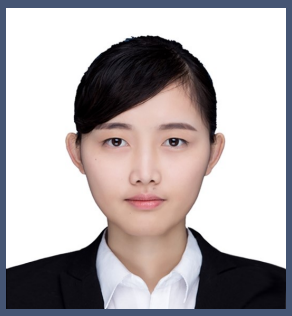
Shan **Gao**



Profile Summary

Experienced **Biostatistician** with 5+ years of international expertise in **clinical trials** and **real-world data**. Proficient in **SAS** and **R** with strong background in **survival analysis** and **predictive modeling**. Skilled in cross-functional collaboration and regulatory-compliant reporting.

Contact Details

@ gaoshan442475397@gmail.com
+32 484 7008 62
in linkedin.com/in/shan-gao-stats

Skills

- SAS (BASE, STAT, GRAPH)
- R (tidyverse, survival)
- Microsoft Excel, MS Word,MS PowerPoint
- HPC
- GitHub
- ICH-GCP
- CDISC (SDTM, ADaM)
- Communication and team collaboration

Languages

Chinese (native)
English (fluent)
Japanese (conversational)
Dutch (conversational)

Hobbies

Photography
Gardening
Hiking

Experience

PhD researcher **2021.02–2025.04**
KU Leuven | Leuven, Belgium

◇ Managed the cleaning and preprocessing of *large-scale EHR datasets* spanning over 8 years (2012–2020) from UZ Leuven, using reproducible R workflows to ensure data quality for modeling, enabling downstream *risk prediction research*.

◇ Collaborated with nurses, clinicians, and infection prevention specialists to standardize clinical data definitions and select relevant variables, enhancing the *clinical relevance* and adoption potential of predictive models.

◇ Implemented and compared multiple modeling strategies on *high-performance computing clusters*, ensuring robustness and scalability across longitudinal EHR data.

◇ Developed and validated dynamic supermodels using the *landmarking approach* to predict CLABSI risk in real-time hospital settings, enabling timely identification of high-risk patients and supporting tailored interventions.

Biostatistician **2019.04–2021.02**
Astellas Pharmaceutical Inc. | Tokyo, Japan

◇ Provided statistical input on *Phase I–III study* protocols, including *sample size calculations* and *statistical analysis plan* design, ensuring *regulatory compliance* and methodological rigor.

◇ Drafted statistical sections for *protocols* and *Clinical Study Reports*, supporting submissions to global regulatory agencies (*PMDA*, *FDA*).

◇ Collaborated with statistical programmers to design *SDTM* and *ADaM* datasets aligned with *CDISC* standards, producing precise specifications for *TLFs* to ensure compliance and reproducibility.

◇ Worked with data managers and clinical investigators to harmonize *SAPs* with evolving data collection, minimizing inconsistencies between planned and actual analyses.

◇ Conducted *interim and final analyses* using *SAS*, delivering accurate and timely outputs for regulatory submission.

◇ Participated in routine *cross-functional meetings*, providing statistical perspectives during protocol reviews, data reviews, and safety updates, contributing to efficient team decision-making.

Education

Doctor of Biomedical Science *KU Leuven* **2021–2025**

Master of Statistics *KU Leuven* **2016–2018**

Bachelor of Chemical Biology & Double Degree Bachelor of Economics (Mathematical Statistics) *Xiamen University* **2012–2016**

Publications

Full list available at: ORCID 0000-0003-0871-1694

◇ Gao S, Albu E, Putter H, et al. A comparison of modeling approaches for static and dynamic prediction of central-line bloodstream infections using electronic health records (part 1): regression models. *Diagnostic and Prognostic Research* 9(1), 20 (2025).

◇ Gao S, Albu E, Tuand K, et al. Systematic review finds risk of bias and applicability concerns for models predicting central line-associated bloodstream infection. *Journal of Clinical Epidemiology* 161, 127–139 (2023).