Jaime Yan

Morristown, NJ 07960 (475) 218-8823 yanmingyunmt@gmail.com

EXPERIENCE

Merck — **Rahway, NJ**
Associate Principal Scientist, Statistical Programming — Nov 2021 – Present

- **Strategic Leadership & Governance:** Provide strategic leadership for statistical programming across multiple concurrent Phase 1-3 cardiovascular studies; define programming strategies aligned with study objectives, timelines, and corporate standards; collaborate with Biostatistics leadership on resource allocation and milestone planning; contribute to development of Statistical Programming SOPs, work instructions, and reusable code frameworks.
- **Study Delivery & Technical Oversight:** Oversee end-to-end statistical programming deliverables including SDTM, ADaM, TLFs, and integrated analyses; lead NDA/BLA eSubmission packages with Define.xml v2.1, reviewer's guides, eCTD, and ADRG in 21 CFR Part 11 validated environment; ensure CDISC compliance and Pinnacle 21 validation; support database lock, interim/final analyses, and regulatory inspection readiness.
- **Innovation, Automation & AI:** Lead GenAI automation initiatives; architected enterprise AI platform with LLM/RAG pipeline enabling programming productivity gains; created **py4csr** (PyPI), an open-source Python framework for automated clinical study reporting; drive adoption of automation, modular programming, and AI-assisted solutions with appropriate validation and regulatory compliance governance.
- **Vendor Oversight:** Provide oversight of CRO programming vendors including evaluation of technical capabilities and automation maturity; ensure vendor compliance with internal standards and expectations; identify and mitigate risks related to timelines and quality; establish data transfer specifications for external vendor data integration.
- **Cross-Functional Collaboration:** Serve as statistical programming representative on Data Standards Committee and cross-functional study teams; partner with Biostatistics, Data Management, Clinical, and Regulatory stakeholders; participate as subject-matter expert supporting audit and inspection readiness with full traceability and documentation.
- **Advanced Analytics & R/Python:** Support use of R and Python for exploratory analyses, data visualization, and advanced analytics; built AI-powered code translation tool for SAS-to-R migration; developed R Shiny applications for dynamic data review; conducted statistical simulations (MCMC, Bayesian) supporting clinical decision-making.
- **Talent Development:** Mentor programmers on CDISC standards, modern programming practices, and AI/automation tools; contribute to recruitment; active PHUSE/PharmaSUG contributor promoting innovation best practices.

Regeneron — **Basking Ridge, NJ**
Senior Statistical Programmer — Jun 2020 – Nov 2021

- **Clinical Programming & Delivery:** Developed SDTM and ADaM datasets for Phase 1-3 oncology trials including solid tumors and dose-escalation studies; implemented RECIST 1.1 tumor response endpoints; created ADPC/ADPP

datasets for PK sub-studies; collaborated with Biostatistics on SAP implementation and Data Management on specifications.

- **Rapid Delivery & Vendor Integration:** Delivered ad-hoc analyses for DSMB meetings and regulatory interactions; managed external vendor data integration ensuring submission-ready quality and CDISC compliance.

Boehringer Ingelheim **Ridgefield, CT**
Statistical Programmer Feb 2018 – Jun 2020

- **Integrated Analyses & Standards:** Developed ISE/ISS pooled datasets for oncology and cardiovascular regulatory submissions; designed reusable SAS macro libraries improving efficiency and standardization; contributed to TAUG development following CDISC standards.
- **Automation & Quality:** Contributed to DaVinci open-source initiative (R Shiny visualization); established automated Pinnacle21 compliance workflows; supported quality control procedures aligned with regulatory requirements.

Talentech **Monmouth Junction, NJ**
SAS Programmer May 2017 – Feb 2018
Built foundation in clinical trial programming: generated TLFs using standard ADaM datasets (ADSL, ADLB, ADAE, ADTTE, ADRS) with strict CDISC compliance; developed core SAS macro expertise.

CORE COMPETENCIES

- **Technical Expertise:** SAS/SAS Macros (Expert, Certified), R (analysis, visualization, Shiny), Python (automation, LangChain, RAG), SQL, Linux, Git; statistical computing environment (SCE) optimization
- **CDISC & Submissions:** SDTM, ADaM, TLFs, Define-XML v2.1, Reviewer's Guides, Pinnacle21, ADRG/SDRG; FDA/EMA submissions (NDA/BLA), eCTD packages, 21 CFR Part 11, ICH E6(R2)/E9/GCP compliance
- **Innovation & AI:** GenAI/LLM platforms, automation frameworks, AI-assisted programming solutions with validation governance; open-source contributions (py4csr, DaVinci)
- **Leadership:** Strategic programming leadership, CRO/vendor oversight in outsourced models, cross-functional collaboration, SOP development, talent mentorship
- **Therapeutic Areas:** Cardiovascular, Oncology (RECIST 1.1), Clinical Pharmacology/PK; Phase 1-3 clinical development

EDUCATION

Harrisburg University of Science and Technology *PA, US*
PhD, Data Science (Expected 2026) *Aug 2022 – Present*
Research: AI/ML applications in clinical development with focus on regulatory compliance, validation, and automation governance

New Mexico Tech *NM, US*
Master of Science, Applied Mathematics & Engineering *Jun 2017*

Jilin University *Changchun, China*
Bachelor of Science *Jun 2015*

CERTIFICATIONS

- SAS Certified Advanced Programmer for SAS 9
- SAS Certified Base Programmer for SAS 9