**David Michael Pressley**

Technologist in the Life Sciences Industry

#### [david@clearcreekanalytics.com](mailto:david@clearcreekanalytics.com)

I am a lifelong learner with 20 years of experience in the Life Sciences spanning medical research, drug discovery and drug development. I provide competent leadership and individual contribution for biostatistical workflows in the drug development value stream. My expertise is in configuring and enabling analytical systems and ETL workflows and artifacts (primarily using SAS) from protocol development to regulatory post-submission. Engaging users, adapting user requirements, and promoting novel solutions is a soft skill I continually work to improve. My ideal future state is to work with like minded, kind professionals in a high trust, collaborative environment using modern technologies and methods.

EXPERIENCE

Roivant Sciences Ltd— ***Digital Innovator***

January 2018 - July 2018

Provided leadership, subject matter expertise, and individual contribution for the configuration and enablement of systems of record within the Biostatistics, Clinical Pharmacology, and Nonclinical functional areas. Effectively served as a systems architect. Systems, applications and responsibilities include:

* Red Hat Linux 7 terminal server (VirtualBox and AWS configured with ssh and GUI access)
  + Base SAS
  + NONMEM, Perl Speaks NONMEM, Pirana (Clinical Pharmacology Analysis)
  + R/RStudio (Statistical, ClinPharm and nonclinical Analysis)
  + Git and GitLab (source code and workflow management)
* Windows Server 2016 (jump server and Windows specific application server)
  + WinNonLin (NCA/PK/PD/TK modeling)
  + PASS (Sample size calculator)
  + MobaXterm(terminal application with X Windows system for access to Linux SAS)
* Creation of bash scripts to:
  + Configure the user’s local working area for project based biostatistical workflows
  + Generate ssh keys for secure access to git and GitHub repositories over ssh
  + Configure Git LFS for efficiently versioning binary files such as SAS datasets
  + Set up synchronization (rsync) and scheduling (cron) of data from vendor SFTP site to user’s local project area
* Utilization of git, GitHub, GitLab, and Git Large file storage to ensure traceability of the biostatistical programming workflow (code, data, artifacts) and system configuration scripts
* Effectively communicated with vendors, colleagues, and management through the use of applications such as Confluence ([HowTo](https://clearcreek.atlassian.net/wiki/spaces/~363221148/pages/596279300/Initializing+a+Project+in+the+Biostatistical+Programming+Workstream?atlOrigin=eyJpIjoiNDA1MDlmYTM0NDMyNDk3N2IwZGYzMmI3YmZiZWEwNzciLCJwIjoiYyJ9), [troubleshooting](https://clearcreek.atlassian.net/wiki/spaces/~363221148/pages/596279300/Initializing+a+Project+in+the+Biostatistical+Programming+Workstream?atlOrigin=eyJpIjoiNDA1MDlmYTM0NDMyNDk3N2IwZGYzMmI3YmZiZWEwNzciLCJwIjoiYyJ9), requirements, knowledge base), Asana (team and project oriented tasks and kanban workflow visualization), Slack (a better alternative to email), and Zoom (video conferencing)
* Negotiated with internal Quality group and external Quality vendors to challenge de facto waterfall approaches to Quality Management
* Responsible for creating and managing vendor creation of Quality documentation surrounding the Red Hat based SAS system
* Collected user requirements, engaged and directed third party vendors for implementation of infrastructure, and application configuration and enablement
* Contributed to the creation of a months long business case for SAS and clinical pharmacology platforms to ascertain the value of each functional group’s application stack as a shared services model
* Led efforts to engage functional areas to create an inventory of all systems in the clinical development value stream. Classifications include:
  + System classification (porcelain or plumbing, system of record or system of engagement)
  + Current status/future status (need/want)
  + Contract model (managed by internal or external)
  + SaaS offering available
  + Quality risk
  + Technology Infrastructure lift required
  + Priority rank

**Clear Creek Analytics, LLC** — ***Contractor and Consultant***

April 2017 - Jan 2018; June 2018 - Current

##### Clinical Data Programmer, Biomarker Data, Gilead Sciences, Inc. (Functional Service Provider role through Triangle Biostatistics, LLC)

* Responsible for generating ETL deliverables for biomarker lab data across multiple therapeutic areas and phases I - III
* Creation and maintenance of programs and automation scripts in a Linux based programming environment

##### Consultant to MedImmune (through d-Wise Technologies)

* Responsible for the creation and maintenance of a custom framework including JMP Clinical, SAS 9.4 and Apache Velocity template to address SDTM data for the creation of custom AE Narratives

##### Consultant to Triangle Biostatistics

* Provided expertise in a risk based, fit for purpose validation of Triangle Biostatistics hosted SAS programming platform
* Utilized Atlassian Confluence for creation of validation documentation, backup and recovery, and disaster recovery SOPs
* Documented user requirements and created validation documentation traceable to each requirement. This included risk assessment, validation plan, IQ/OQ and acceptance testing documentation. Executed tests and signed off on approvals

**United Therapeutics Corporation** — ***Associate Director, Statistical Programming and Analytical Systems Architect***

March 2006 - March 2017 (Started with title Clinical Data Programmer)

My responsibility as the Associate Director was to ensure the needs of our customers were met through accurate representation of data across the clinical development lifecycle. I led the design and implementation of our analytical environment which utilized virtualization and open source technologies such as Git, Subversion, Python and R.

I provided individual contribution and leadership of the statistical programming group supporting clinical trials of all phases in two major therapeutic areas: pulmonary arterial hypertension and pediatric neuroblastoma, and provided representation of the biometrics group at internal meetings and regulatory agency meetings.

PROJECTS OF SIGNIFICANCE

### 2015: Remodulin Implantable System Supplemental New Drug Application

Responsible for the preparation of regulatory filing documents and data in formats considered to be reviewable by FDA’s Center for Drug Evaluation and Research (CDER) division. This was considered a joint review by the medical device division (CDRH) and CDER. Each division had vastly different expectations surrounding data standards and formats. I was initially brought on to advise and review, but at the last minute designed and wrote a significant amount of code to transform the data into an acceptable format.

### 2012 - 2014: Unituxin Biological Licensing Application and Approval

Led the Statistical Programming and Data Management team for Unituxin BLA Submission to FDA and MAA submission to EMA. This was a complicated cooperative group project that allowed me to learn how to deal with data the organization had minimal control over. Data structures were very denormalized, and the state of data was constantly changing due to the data transfer schedule and unlocked database. Without an intentional design to ensure traceability and lineage of data and code, this project would not have been a success.

**2006 - 2012: *Orenitram New Drug Application and Approval***

Principal Programmer and leader of the Statistical Programming team for Orenitram NDA Submission to FDA. This project encompassed three, Phase 3 trials and twenty one Phase 1 and 2 supporting trials, along with the required data formats, data specifications, and relevant documentation

**2007 - 2009: *Tyvaso New Drug Application and Approval***

Individual Contributor for Statistical Programming and Data Management support for Tyvaso NDA Submission to FDA. This project was challenging because a team of 2 performed both the data management edit-check and data verification programming; statistical programming to show drug efficacy and safety; and data programming to meet regulatory agency guidelines and regulations for submission.

BRIEF SUMMARY OF PROFESSIONAL EXPERIENCE:

* 11+ years as a clinical data and statistical programmer. 8 years experience creating and utilizing CDISC data.
* Thorough understanding and implementation of statistical programming lifecycle including protocol review, CDMS specification and edit check logic, creation of programming specifications, metadata driven creation of SDTM and ADaM domains, conformance checks, submission of data definition documentation, reviewer’s guides and XPTs, addressing agency RFIs and representation at regulatory agency meetings
* 4 regulatory agency submissions and approvals (3 FDA and 1 EMA)
* Individual contributions related to data standards include: responsibility for metadata specification, implementation of required CDISC compliant datasets, tables, listings and figures, and regulatory filing documentation across all phases of clinical development lifecycle, including ISS/ISE
* Administrator of statistical computing platform. Responsible for the development, validation and implementation of the Statistical Computing Environment used to perform the business function of the biostatistics team
* Creation of data safety monitoring board output and interim reports
* Creation of planned and ad-hoc output to support clinical study reports
* Interaction with Principal Investigators for the creation of publication quality graphics using scalable vector graphics and SAS/GRAPH
* Collaboration with Data Management colleagues to provide standardized input into design of data capture (CDASH)
* Negotiation and administration of software licenses, service provider contracts, and server hosting contracts for the biometrics group
* I am a proponent of the adoption of standard software development practices and the utilization of open source technologies to ensure the quality of data and programming workflows. Graph databases and resource description frameworks are becoming increasingly popular, and can potentially enable the automation of clinical workflows; extracting metadata from protocol specifics, transformation to SDTM and ADaM structures, and linking to agency submission deliverables. This persistence layer could potentially be leveraged for machine learning and optimization of clinical development workflows.
* Linux and OS virtualization, along with cloud based technologies are becoming increasingly popular, and can potentially lower the time barrier for creation of infrastructure solutions to address business needs.

AREAS OF TECHNICAL INTEREST : SKILL LEVEL / [EXAMPLE CODE]

* Python : Intermediate / [[readFile.py](https://gist.github.com/pressleydavid/939e6871189c6a13c7ca), [readXPT.py](https://gist.github.com/pressleydavid/733eba5b5094d58c4a4e), [rm\_bookMarks.py](https://gist.github.com/pressleydavid/fa8c6437ea9818361d9d), [createFolderTree.py](https://gist.github.com/pressleydavid/98c9d46d2e983aae6969), [sas7bdat.py](https://gist.github.com/pressleydavid/15f5d55d8e576f035bb3)]
* SAS Foundation/Graph/Stat/Macro/etc modules : Expert / [[readSpec.sas](https://gist.github.com/pressleydavid/5367151583f9a0b29c2c0edb47d7b542), [hashtable.sas](https://gist.github.com/pressleydavid/2fb3e93467dfd652a2507949afd5766b), [addLabels.sas](https://gist.github.com/pressleydavid/4446ee586068bca0e5f5), [createXPT.sas](https://gist.github.com/pressleydavid/f80ca5146a2fb9419127), [unzipXPTLibrary.sas](https://gist.github.com/pressleydavid/1c570253c086f32ee788), [histogramShifts.sas](https://gist.github.com/pressleydavid/eed1aa17b428a210620e), [hemodynamicsGraphs.sas](https://gist.github.com/pressleydavid/ae7ca2cbc20f9f660964), [init.sas](https://gist.github.com/pressleydavid/22bd3ce58ef96d95404b0be4ae08a6eb), [getProjectPath.sas](https://gist.github.com/pressleydavid/1e59ac4936fb5841e1cfa061d39f4ea2), [getdatasetnames.sas](https://gist.github.com/pressleydavid/d5169a590657d1e379736db4d03e77f5)]
* Subversion : Expert
* Git : Intermediate: understand the basic concepts and commands of common workflows and branching between local user area and remotes; some use of ‘plumbing’ commands in scripting
* R : Academic
* C++ : Academic
* Apache Spark: Academic
* Neo4j : Learning
* Docker : Learning
* HTML/CSS: Academic
* JavaScript: Academic
* Unix/Linux: Intermediate. Continuously learning, but reasonably well versed in bash commands and scripting for system administration purposes
* Vagrant: Intermediate. Adept at provisioning virtual development servers using Vagrant with shell scripting to provision common applications and settings.
  + RedHat 7 with a duplication of the analytical toolset used in production (SAS, Python, git, GitLab, UltraEdit, MobaXterm, TigerVNC server for enabling desktop connections).
  + Ubuntu 16.04 with Neo4j, Apache web server, virtualenv, django, lxml, mod\_wsgi, and openpyxl to surface a running Neo4j instance on localhost

PREVIOUS WORK HISTORY

**Amphora Discovery** — ***Programmer/Analyst***

September 2001 - March 2006

Started as Bench Scientist to support High Throughput Screening using an emergent microfluidic technology (Caliper Life Sciences) used to screen relevant signal transduction pathways (mostly kinase and phosphatase targets) against libraries of significant size. We generated about 1 Million data points per week.

##### Responsibilities as a Programmer/Analyst:

Database registration and record maintenance of all small molecule compounds received by the company

Responsible for creation of IDBS ActivityBase database templates and ad-hoc Excel VBA and SQL tools to enable access to information

**Glaxo Smith Kline** — ***Research Associate (contractor)***

August 2001 - January 2002

Bench work contributing to development of drug candidates for the treatment of cancer

**Nortel Networks** — ***Equipment Applications Engineer***

October 2000 - August 2001

Hardware engineer responsible for provisioning and configuring public and private fiber optic network switch systems

**Duke University Medical Center** — ***Research Technician***

December 1998 - October 2000

Bench Scientist in the lab of Dr. YT Chen executing basic research of Phases I / II clinical trials to test efficacy of enzyme replacement therapy in infantile and juvenile Glycogen Storage Disease Type II (Pompe).

EDUCATION

## Coursera

29 November 2016

# [HTML, CSS, and Javascript for Web Developers](https://www.coursera.org/account/accomplishments/certificate/LVPFSDYRM8V3)

## EdX

10 July 2015

# [CS100.1x: Introduction to Big Data with Apache Spark](https://s3.amazonaws.com/verify.edx.org/downloads/446c9a2aba924794b17e3e4f3263fd5e/Certificate.pdf)

6 August 2015

# [CS190.1x: Scalable Machine Learning](https://s3.amazonaws.com/verify.edx.org/downloads/0be48224c88540999ed6344243b6a847/Certificate.pdf)

4 September 2015

# [DAT204x: Introduction to R Programming](https://s3.amazonaws.com/verify.edx.org/downloads/eb014507243c4eca90c52e63c63bc911/Certificate.pdf)

6 November 2015

# [6.00.1x: Introduction to Computer Science and Programming Using Python](https://courses.edx.org/certificates/8c77c1a8a4df4d5abed5fe72c0547fef)

**North Carolina State University** — ***Graduate Coursework in Statistics***

Fall Semester 2008 - Spring Semester 2010

# Fall 2008: ST 511 - Experimental Statistics For Biological Sciences

# Spring 2009:ST 512 - Experimental Statistics For Biological Sciences

# Spring 2010: ST505 - Applied Nonparametric Statistics

**North Carolina State University** — ***Computer Programming Certificate***

Fall Semester 2004 - Spring Semester 2007

**North Carolina State University** — ***Bachelor of Science in Biochemistry / Bachelor of Arts in Chemistry***

Fall Semester 1993 - Summer Term I 1998

PUBLICATIONS AND PRESENTATIONS

Wu G, Irvine J, Luft C, Pressley D, Hodge CN, Janzen B. Assay development and high-throughput screening of caspases in microfluidic format. Combinatorial Chemistry and High Throughput Screening, 2003 Jun;6(4):303-12

McVie-Wylie AJ, Ding EY, Lawson T, Serra D, Migone FK, Pressley D, Mizutani M, Kikuchi T, Chen YT, Amalfitano A. Multiple muscles in the AMD quail can be "cross-corrected" of pathologic glycogen accumulation after intravenous injection of an [E1-, polymerase-] adenovirus vector encoding human acid-alpha-glucosidase. Journal of Gene Medicine, 2003 May;5(5):399-406

Ding EY, Hodges BL, Hu H, McVie-Wylie AJ, Serra D, Migone FK, Pressley D, Chen YT, Amalfitano A. Long-term efficacy after [E1-, polymerase-] adenovirus-mediated transfer of human acid-alpha-glucosidase gene into glycogen storage disease type II knockout mice. Hum Gene Ther. 2001 May 20;12(8):955-65.