132

Navigating the Move to Open Source - Effective Strategies for Adoption and Working with Different Software (AIMS SIG and CAMIS)

Mark Bynens[1], Martin Brown[2], Lyn Taylor[3], Christina Fillmore[4], Agnieszka Tomczyk[5]
[1]J&J, Brussels, Belgium. [2]PPD, part of Thermo Fisher Scientific, Cambridge, United Kingdom. [3]Parexel, Sheffield, United Kingdom. [4]GSK, London, United Kingdom. [5]Parexel, Warsaw, Poland

Mark Bynens
Please provide a brief biography for the Presenting author(s)
Mark Bynens is Director and Scientific Computing Operations (SCO) Head within Statistics & Decision Sciences (SDS), Global Development, J&J innovative medicine. In this role, he, together with his team, is responsible and accountable for amongst other: change management, project management and end-to-end management of software applications for statistical evaluation or for business processes, education in-classroom and e-learning, knowledge sharing, software/application acquisition and high-performance computing for intensive data evaluation, simulations, and statistical research. He is one of the main authors of the SCE White Paper. Mark has more than 15 years of experience in the industry and joined Janssen as a full-time employee in 2016. He studied at Catholic University of Leuven and earned a master’s degree in mathematics.

Martin Brown
Please provide a brief biography for the Presenting author(s)
Martin Brown currently works as an Associate Director in PPD’s Statistical Science group having joined PPD in 2007. In this role he consults on many different areas including adaptive designs and simulations (including CRM and basket designs), statistical software, centralized statistical monitoring and is leading the move towards using R. He received his MSc. in Statistics from the University of Sheffield and is a Chartered Statistician with the RSS. Martin is the chair of the PSI AIMS SIG and is also a member of the Visualisation SIG.

Lyn Taylor
Please provide a brief biography for the Presenting author(s)
Lyn Taylor has worked in medical research for over 23 years mostly working in phase I-III clinical trials in a wide range of therapeutic areas. After 11 years at PAREXEL, Lyn gained experience at PRA Health Sciences, PHASTAR, and in 2022 moved back to PAREXEL as an Associate Director Biostatistics. In 2014, Lyn obtained her Ph.D. from Sheffield University in Statistical Modelling of Markers of Severity in Rheumatoid Arthritis. Lyn is co-lead of CAMIS and has previously been an active contributor to the PSI AIMS SIG and R Validation Hub. Her current interests are in documenting the differences in the methodology used by SAS and R to allow concurrent use of the software systems.

Christina Fillmore
Please provide a brief biography for the Presenting author(s)
Christina Fillmore formally trained as a statistician, but over her 8 years at GSK transitioned into a data scientist. With a keen interest in R, she's co-lead of the PHUSE CAMIS project and R/pharma Diversity Alliance. She maintains metacore, metatools, and beastt R packages. Her focus is on creating open-source packages that enable others to deliver studies faster.

Agnieszka Tomczyk
Please provide a brief biography for the Presenting author(s)
Agnieszka Tomczyk is a Senior Biostatistician in Parexel, has graduated from Applied Mathematics from the Technical University of Lodz and has been working in medical research less than 5 years. Agnieszka has previously worked in R as a data scientist, now has been programming in SAS® as well for the last couple of years. She is interested in

comparing the functions, methods and packages in R, SAS and other statistical software. Recently joined the PHUSE/PSI CAMIS project building the open source github repository.

Single topic, multi-speaker session, Workshop or Single presentation submission

A single topic, mutli-speaker session/workshop

Single topic session or workshop abstracts

Talk 1 - Mastering the Art of Adopting R and Python: Innovative Strategies for Effective Change Management

Author - Mark Bynens

Abstract: Mastering the Art of Adopting R and Python: Innovative Strategies for Effective Change Management.' is more than just a presentation; it's a roadmap to navigate the complexities to integrate R and Python into our daily operations in a world that never slows down. Through an in-depth look at some real-world examples moving towards R and Python we will show you how it's done. This isn't just theory; it's practical, actionable advice. As we embark on a journey to weave R and Python into the fabric of our organization, let's keep these insights and strategies at the forefront. Together, we can redefine what it means to be adaptable and resilient in an ever-changing world.

Talk 2 - R Adoption & Change Management – a Large CRO Perspective

Author - Martin Brown

Abstract: In today's rapidly evolving technological landscape, organizations are increasingly leveraging open-source tools to enhance their statistical analysis, reporting, competitiveness and agility. For Contract Research Organizations (CROs), this transition can be challenging, as much of the current work is traditionally conducted in SAS, supported by well-established and efficient processes. This presentation will address the complexities and resource constraints associated with such a transition. Attendees will gain valuable insights into best practices for implementing change management strategies, enriched by a detailed case study of how change management is being implemented in a large CRO Biostatistics and Programming department. This case study will demonstrate how R and other open-source tools can drive efficiency and foster innovation right now while moving as fast as possible towards a future state. By incorporating key change management aspects—such as leadership, vision, training, competency, support and communication—viewed through the lens of a large CRO context, this presentation will equip you with essential knowledge and considerations for achieving seamless and effective change management within your organization. Whether you are part of a large or small Pharma, Biotech, or CRO, this presentation will provide you with thought-provoking insights for a successful transition to R and open-source solutions.

Talk 3 - R you (all) right, SAS? – Replicating statistical results between software

Authors - Lyn Taylor, Christina Fillmore & Agnieszka Tomczyk

Abstract:

Comparing Analysis Method Implementations in Software (CAMIS) is an open source repository to document differences in statistical methodology across software. Statisticians and programmers often work in multiple software systems (e.g. SAS, R, Python, StatXact, EAST) and may experience inconsistencies in analysis results. Our highly regulated medical research industry, generally uses double (or triple!) programming to validate the analysis, which requires an identical match in results.

Firstly, we will showcase the PHUSE CAMIS project idea. The project investigates differences and similarities between various software and stores all code, case studies, results, differences and findings in an open source github repository (https://psiaims.github.io/CAMIS/). We will give the common reasons for discrepancies and explain how to avoid them.

Secondly, we will present a few project examples to demonstrate the sorts of discrepancies we are observing. That will include survival analysis, confidence intervals for proportions, non-parametric methods and sample size calculations.

With the current shift in the industry from using SAS to R and other open source software we want to draw attention to the importance of CAMIS project. The audience will leave with the tools they need, to efficiently investigate discrepancies between software, that they experience in their day to day work. We will clear up any concerns and reassure the current and future R users that in the majority of cases, the results can be replicated across software. Where results cannot be replicated, the CAMIS repo explains the reasons why.