PHUSE EU Connect 2023
Paper LI09
Stream: Leadership Through Innovation (LI)

Becoming an Integrated Part of the Open-Source Pharma Community

Phil Bowsher, Posit/RStudio PBC; Boston, USA

Abstract

Many pharmaceutical organizations have migrated or are shifting to an open-source backbone in clinical trials. The video below provides information on this shift for Roche:

Roche: Shifting to an Open-source Backbone in Clinical Trials
https://www.youtube.com/watch?v=nqJsLSLd39A

Novo Nordisk's Journey to an R based FDA Submission
https://www.youtube.com/watch?v=t33dS17QHuA

A list of organizations using open-source in submissions can be found here:

https://github.com/philbowsher/Open-Source-in-New-Drug-Applications-NDAs-FDA

As new organizations journey into the open-source space, they often wonder:

“How does my pharma get involved in Open Source Drug Development?”

A main motivation beyond contributing back to the open source space, is **learning what other pharmas are doing**. Open source has led to an explosion in cross-collaboration in pharma that is bringing significant efficiency gains in clinical reporting such as time savings in creating TLGs. This collaboration helps pharmas in many ways such as: getting the latest developments more rapidly, helping with hiring new talent, switching between languages, and collaborating with external partners. Many of these efforts have been focused on **"pan-pharma" code collaboration** for Clinical Reporting.

There are a lot of great open-source programming languages used for drug development such as R, Python, Julia, Javascript, SQL, Stan etc. Version control has opened the door to collaboration. Below, I will review the current ecosystem for open-source drug development and highlight boundaries broken via open-source collaboration.

Pan-Pharma Cross-Collaboration

This talk will highlight current success stories and discuss the new horizon for new drug submissions. I will also feature ways that you and your pharma can get involved in this exciting new effort that is bringing pharma together!

Open-source drug development is transitioning the clinical reporting space as pharmas are collaborating on the code used to create the workflows and reports. J&J, Roche, GSK and Novo Nordisk are leading the charge on incorporating open-source into drug development.

This change is common in other industries such as technology where groups commonly collaborate on technology stacks. Common examples include Spark, Docker etc. Thomas Neitmann at PHUSE US Connect 2023 has a great quote about this”

"**We compete on the molecules, not the software**…because open-source software is often developed collaboratively, it can benefit from a faster pace of innovation than proprietary software. New features and capabilities can be added more quickly, and bugs and issues can be addressed more efficiently."

https://phuse.s3.eu-central-1.amazonaws.com/Archive/2023/Connect/US/Florida/PRE_OS11.pdf

There have other excellent talks recently about cross-collaboration in pharma around open-source such as:

Ross Farrugia - "Breaking boundaries through open-source collaboration"

https://www.youtube.com/watch?v=APMDOS4v9Hk

Roche has a site dedicated to cataloging Roche Pharma Data & Statistical Sciences open source activity:

https://codecollaboration.org/opensource/

Return on Investment

As pharmaceutical organizations embark on op-source drug development, a first question usually focuses on the Return on Investment (ROI). A recent talk by Roche highlights many gains such as:

https://insightsengineering.github.io/nest-2023-summeR/resources/NEST_%20Now%20&%20Future.pdf

- Partnerships and collaborations with external partners
- Streamlined Workflows
- Talent out of university more likely to know open-source programming languages
- latest developments more rapidly
- switch between languages and contexts more easily
- time savings in creating standard TLGs
- time savings in creating customized TLGs
- access to free solutions to support creating a clinical submission

Where to Help & How to Get Involved

Pharmaceutical organizations often wonder how to get involved in Open Source Drug Development. There are great open-source efforts happening currently. Where to help depends on your company's focus and where to add the help. First we will review some of the free initiates that require no membership or fee and help provide information and networking:

Current Free Pharma Organizations and Efforts:

1. **Open Source in Pharma & R in Pharma - rinpharma.com**

Open Source in Pharma is the 501(c)(3) that manages R in Pharma and other pharma related efforts listed here:

https://opensourceinpharma.com/

R in Pharma was started in 2017 by many pharmaceutical companies that wanted a conference focused on open-source drug development. Today R in Pharma has grown into a community effort that has over 5K registrations. R in Pharma was born out of Harvard University and now has 2 main events: Virtual Gathering and Leadership Summit.

The main focus of R in Pharma is "Community and learning".

R/Pharma has various committees:

- Program Committee - this group helps put together the talks and content at the conference.
- Organizing Committee - this group helps run the conference and the members are here:

https://rinpharma.com/people/

You can reach out to R in Pharma to help on a committee here:

https://rinpharma.com/contact/

2. **R Validation Hub - pharmar.org**

R Validation Hub is a collaboration to support the adoption of R within a biopharmaceutical regulatory setting. The focus is on the “Validation” topic and the popular "A RISK-BASED APPROACH FOR ASSESSING R PACKAGE ACCURACY WITHIN A VALIDATED INFRASTRUCTURE" white paper was written by GSK, Janssen & Pfizer. Last year, R Validation Hub won "Best App" at Shiny Conf for the {riskassessment} Shiny app.
Demo app:
http://rinpharma.shinyapps.io/riskassessment/

Video: {riskassessment} App for R-Package Validation
http://youtube.com/watch?v=gsWc_oSTb9c

GitHub:
http://github.com/pharmaR/riskassessment

You can join and get involved here:

https://www.pharmar.org/contribute/

https://www.pharmar.org/contact/

3. *Regulatory R Package Repository*

Validation is a popular topic in pharma. A sub-group of the R Validation Hub is the Regulatory R Package Repository. This group is focused on a cross-industry approach of establishing and maintaining a 'repository' of R packages with accompanying evidence of their quality and the assessment criteria, that can be used to simplify necessary in-house validation processes as much as possible. You can join and help here:

https://github.com/pharmaR/regulatory-r-repo-wg/issues/1

https://pharmar.github.io/regulatory-r-repo-wg/#how-to-get-involved

4. *Openstatsware*

The American Statistical Association (ASA) Biopharmaceutical Section (BIOP) Software Engineering Working Group (SWE WG) is focused on engineering selected R-packages, developing good SWE practices for engineering high-quality statistical software and communicating and collaboration. Information on join is here:

https://community.amstat.org/biop/workinggroups/swe-wg

https://rconsortium.github.io/asa-biop-swe-wg/join_us.html

This group teaches a popular workshop called “GOOD SOFTWARE ENGINEERING PRACTICE FOR R PACKAGES”:

https://openpharma.github.io/workshop-r-swe-md/

5. *Pharmaverse*

Pharmaverse is a connected network of companies and individuals working to promote collaborative development of curated open source R packages for clinical reporting usage in pharma. A popular package is admiral for a modularized toolbox that enables the pharmaceutical programming community to develop ADaM datasets in R. You can see a list of all the packages here:

https://github.com/pharmaverse/

You can join the Pharmaverse slack here:

https://pharmaverse.slack.com/

Information on contributing is here:

https://pharmaverse.org/contribute/

6. *Bioconductor*

Bioconductor is a project to develop, support, and disseminate free open source software that facilitates rigorous and reproducible analysis of data from current and emerging biological assays.

New Developer Program”

https://www.bioconductor.org/developers/new-developer-program/

Community Advisory Board:

https://bioconductor.org/about/community-advisory-board/

Slack:

https://slack.bioconductor.org/

7. Phuse Working Groups:

PHUSE has various open-source initiatives focused on diverse areas of drug development:

https://advance.phuse.global/display/WEL/Working+Groups

You can contact the various Project leads at the link above about joining the effort.

8. R Submissions Working Group

The R submission working group is a cross industry pharma working group focusing on improving practices of R-based clinical trial regulatory submissions. This group has completed 2 successful pilot submissions to the FDA and have other pilots in the works. You can watch previous meetings and learn more here:

https://rconsortium.github.io/submissions-wg/

You can request to join by submitting an issue here:

https://github.com/RConsortium/submissions-wg

9. R Tables for Regulatory Submissions Working Group

R Tables for Regulatory Submissions (RTRS) includes representation from several large pharmaceutical companies and CRO's. The goal of the working group is to create standards for creating tables that meet the requirements of FDA submission documents, and hence enhance the suitability of R for FDA submissions. This group has created a R Tables book here:

https://rconsortium.github.io/rtrs-wg/

You can request to join by submitting an issue here:

https://github.com/RConsortium/rtrs-wg

10. CAMIS (Comparing Analysis Method Implementations in Software)

CAMIS is a cross-industry group formed of members from PHUSE, PSI, and ASA. The focus of this group is to provide guidance on the types of questions statistical staff should ask to identify the fundamental sources of discrepant results between software. The group welcomes contributions from the wider community and has a great page comparing and documenting differences in analysis method implementations in software.:

https://psiaims.github.io/CAMIS/

Please open an issue here to join or add information to the Github:

https://github.com/PSIAIMS/CAMIS

Other Initiatives:

Below is a list of other initiatives with open-source components:

- psiweb.org
  A community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients. Various working groups can be found here such as the Wonderful Wednesdays brought to you by the PSI Visualisation SIG: https://psiweb.org/sigs-special-interest-groups/sigs
- transceleratebiopharmainc.com
  TransCelerate is a non-profit organization with a mission to collaborate across the biopharmaceutical R&D community.
- www.mbswonline.com
  The Midwest Biopharmaceutical Statistics Workshop fosters discussion on statistical methodology in the pharmaceutical industry.
- hpcxpharma.org
  Mission to (accelerate) improve pharmaceutical research & scientific outcomes by collaborating to share experiences and insights, optimize HPC capabilities, promote standards, and develop best practices.
- go-isop.org
  The mission of the International Society of Pharmacometrics (ISoP) is the promotion and advancement of the discipline of pharmacometrics, through Integration, Innovation, and Impact: quantitative integration of multisource data and knowledge of clinical, biomedical, biological, engineering, statistical, and mathematical concepts, resulting in continuous methodological and technological innovation enhancing scientific understanding and knowledge, which in turn has an impact on discovery, research, development, approval, and utilization of new therapies. Has a conference each year: https://www.go-acop.org/
- page-meeting.org
  PAGE represents a community with a shared interest in data analysis using the population approach. It is a not-for-profit organization (company limited by guarantee) with registered office in England and Wales.
- pharmasug.org
  PharmaSUG is a Software Users Group of life science and health research professionals focused on the application of technological solutions in data analytics and regulatory support.

- https://www.amstat.org/meetings/asa-biopharmaceutical-section-regulatory-industry-statistics-workshop
  The ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop is sponsored by the ASA Biopharmaceutical Section in cooperation with the FDA Statistical Association. The conference lasts two days each year, with invited sessions co-chaired by statisticians from industry, academia, and the FDA.
- https://ropensci.org/
  rOpenSci fosters a culture that values open and reproducible research using shared data and reusable software.
- https://www.cdisc.org/core
  The overall goal of the CORE Project is to deliver a governed set of unambiguous and executable Conformance Rules for each Foundational Standard, and to provide a Reference Implementation of an open-source execution engine for the executable Rules.

Other Ideas for Joining the Open-Source Pharma Community

Designate an Open Source Champion(s) at your Pharma

This person can help share your pharma's open source drug development efforts with the community (liaison) and help be the point person on collaborations. Some organizations have many of these representatives. It is important for the pharmas to dedicate time for these efforts. Below is a small sample of these people at the pharmas listed below:

Harvey Lieberman - Novartis
On the board for Open Source in Pharma and various R in Pharma efforts.

James Black - Roche
On the board for Open Source in Pharma and various R in Pharma and Pharmaverse efforts.

Sumesh Kalappurakal - JnJ
On the board for various R in Pharma and Pharmaverse efforts.

Christina Fillmore - GSK
Data Scientists and R developer supporting various Pharmaverse packages and R in Pharma efforts.

Give Employees Time & Space to Contribute to Open-Source Efforts

'A rising tide lifts all boats'! Your pharma is probably using open-source. Giving back is a great way to say thanks. Employees though need time and space to contribute. It is important to give your employees time for this effort that is not on the weekend or evenings!

Below is a video called "Monitoring health and impact of open-source projects" that provides interesting insight on supporting open-source projects:

https://www.youtube.com/watch?v=kaoe7xuIJ1U

Pharmaverse insights are here:

https://pharmaverse.r-universe.dev/builds

Start a Open Source Drug Development Competency Center at Your Pharma

One of the key items pharmas can do is to start an Open Source Competency Center. This center will be focused on gathering and nurturing the open source drug development community at your pharma. Often this group is led by the

open-source champion(s) listed above. This group can host various events and even a pRodcast! This group will also have a community leader (https://www.youtube.com/watch?v=s0ZU4FDstfY). Ben Arancibia (GSK) shared the experience at GSK on creating “AccelerateR" and how other organizations can use our learnings to scale R from pilots to full enterprise adoption and contribute to open source industry R packages:

https://rinpharma.com/publication/rinpharma_323/

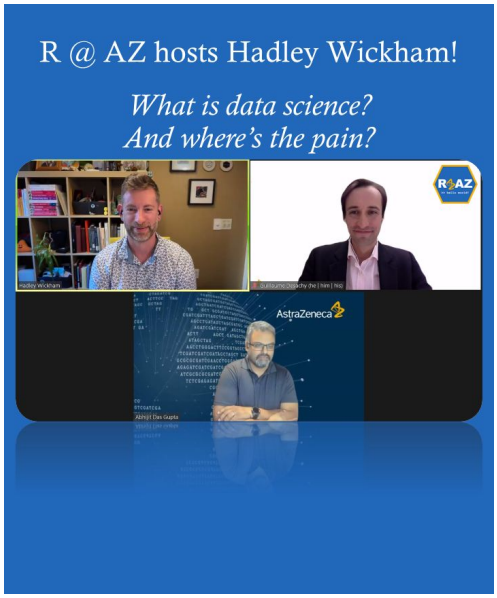
The following presentation has some great information on how Novo Nordisk supports its R users:

https://phuse.s3.eu-central-1.amazonaws.com/Archive/2022/Connect/EU/Belfast/PRE_TT03.pdf

Host an Internal/External Open Source Drug Development Conference

Hosting an internal open-source conference (Virtual or In-Person) is a great way to unite your user communities. These gatherings can help bring awareness to similar projects happening across the company as well as bring insights into where people can get help. Another benefit is that you can invite speakers from other pharmas to join and speak. This is a great way to learn about other efforts and initiatives to join. An example at Astrazeneca:

R@AZ, the community of R users at AstraZeneca, hosted Hadley Wickham on October 19th 2023. This event was Attended by 300+ people and various AZ collaborators!



This could also be hosting an external event like the R/Pharma Summit or a PHUSE single day event.

Start a Internal Slack/Teams Group for Open Source Drug Development

One of the biggest challenges is connecting people at your organization. An internal Slack/Teams Group or channel focused on open-source development can be a great place for people to ask for help or connect with other developers.

Train & Help Employees Become Open-Source Literate

Many of the examples below require a base knowledge in open-source tools and programming. A few key items to train and focus on are:

Reproducibility & Software Engineering CI/CD: Learn about tools/packages like Quarto, Version Control (git), renv, Integrated development environment Software (RStudio, VS Code etc.), Github Actions etc.

Creating & Sharing Code

Below are some ideas on creating and sharing code. This topic is complex and there is excellent guidance below on the use of Open-Source Software (OSS), as well as collaboration on and creation of open-source projects used by data scientists in clinical reporting workflows:

https://phuse-org.github.io/E2E-OS-Guidance/

James Black has a great talk on this subject here:

https://www.youtube.com/watch?v=0PsyMbd9RTA

Dedicate Internal and/or Contractor FTE(s) Working on Fully Open-Sourced Code

Having dedicated people focused on open-sourced code creation can help with the ROI and value from open sourcing code. This is captured often by the collaborative benefits and updates made by others outside your organization back to your code base. An example could be that your organization uses R/Python a lot internally. Creating and sharing a python tool or pipeline could drive adoption at other pharmas. The other group(s) then add contributions back for external use and enhance your tool/package used internally!. This also helps drive the broader open-source development and makes your efforts more acceptable for submissions given the wider adoption. Recruitment then is happy as it is also often easier to recruit people with open-source programming skills thus furthering the innovation internally and externally.

There is a really awesome cross-pharma team that supports admiral:

https://github.com/pharmaverse/admiral/graphs/contributors

Create and Support a Public Github Organization with Packages & Code

Many pharmas have public Github pages dedicated to sharing public packages and code. This is a great way to get additional “pan-company” support as well as share your code with the world. A few examples are:

https://github.com/insightsengineering - Roche

https://github.com/GSK-Biostatistics - GSK

https://github.com/Merck - Merck

https://github.com/Biogen-Inc/ - Biogen

Create, Support and Share Public Package(s)

Many pharmas have public Github, R or Python packages. This could be a Pharmaverse package, an R package on CRAN or a public Python package. Novo Nordisk has a webinar talking about its internal packages here:

https://www.youtube.com/watch?v=t33dS17QHuA

Examples include:

https://nlmixr2.org/ - Novartis with collaborators from various organizations.

https://github.com/GSK-Biostatistics/tfrmt - GSK Pharmaverse Package

Create, Support a Repository/Website of Examples/Articles/Case Studies

Has your pharma created a great Shiny app? Looking for help? A public repository is a great way to share your code and also get contributions back! An example includes:

https://github.com/Biogen-Inc/tidyCDISC - Biogen

https://insightsengineering.github.io/teal/latest-tag/ - Roche

https://gt.rstudio.com/articles/case-study-clinical-tables.html - Richard Iannone, Joe Cheng, Barret Schloerke, Ellis Hughes, Alexandra Lauer, JooYoung Seo

Teach a Public Workshop

Looking for other pharmas to know about your tool/package? Teaching a public free workshop is a great way to create awareness as well as invite others to help! Examples include:

youtube.com/playlist?list=PLMtxz1fUYA5AWYQHB5mZAs-yamNJ5Tm_8 - R/Pharma 2022 Workshops

The code for the public workshops can be shared on github. Users and others then can port the examples directly into RStudio or Posit Cloud. Here is an example:

https://posit.co/blog/creating-adsl-with-the-pharmaverse-part-2/

Here is a ready-to-go environment on Posit Cloud!:

https://posit.cloud/content/5698018

Below is a talk on how JnJ trained legacy stat programming users to learn R is here:

https://github.com/rinpharma/2019_presentations/blob/master/talks_folder/2019-Bargo-Using_RStudio_Cloud_to_advance_R_proficiency.pdf

Give a Talk or Write a Paper on Open-Source

A great way to connect with the open-source drug development community is by giving a talk about your open-source efforts. This could be giving a talk about a package your organization made or how you contributed to a public package.

There are great conferences to showcase this work. R in Pharma, PHUSE, ACOP, ASA BioPharma, PharmaSUG, Bioconductor all have great events that incorporate open-source talks into the conference program.

Conclusion

As interest in open-source increases by clinical statistical programers, pharmaceutical The information outlined in this paper highlights the history of R and Python interoperability and advice when incorporating Python into a clinical space with R.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Phil Bowsher
phil@rstudio.com