Paper SI01

{falcon}: A Collaborative Leap Forward Towards Harmonization of Clinical Reporting Standards

Vincent Shen, Roche, Mississauga, Canada
Juergen Boehl, Boehringer Ingelheim, Biberach, Germany

ABSTRACT

The {falcon} initiative is an industry collaborative effort under {pharmaverse} that unites Boehringer Ingelheim, Moderna, Roche, and Sanofi with the aspiration of building and open-sourcing a comprehensive catalog of harmonized tables, listings, and graphs (TLGs) for clinical study reporting. By leveraging existing open-source R packages, {falcon} aims to simplify the process of output review, comparison, and meta-analyses, while fostering efficient communication among stakeholders in the pharmaceutical industry. In addition, the initiative aligns with CDISC's ARD/ARM effort to promote a standardized approach to data presentation. Drawing inspiration from the FDA Standard Safety Tables and Figures Integrated Guide, the {falcon} Initiative has successfully developed and open-sourced 18 templates to date. The collaborative effort focuses on improving the clarity, consistency, and accessibility of TLGs by addressing variations and redundancies in their creation and use. This harmonized approach allows for streamlined reporting processes and facilitates effective communication of study results within the industry and to regulatory authorities. Future plans for {falcon} involve expanding the catalog through continuous collaboration from participating companies and inviting wider industry engagement. The ultimate goal is to promote harmonization of TLGs for clinical reporting across the pharmaceutical industry, leading to greater efficiency, collaboration, and innovation. In this paper, we will share the collaboration journey of {falcon} so far, providing more details on the current progress, long-term vision, and strategies for this initiative. We will also shed lights on the challenges and opportunities of harmonizing clinical reporting through open-source collaboration and learn about the potential benefits and future direction of {falcon}.

INTRODUCTION

Over the years, pharmaceutical industry has established robust and universally agreed data standards for clinical trial conduct and analyses. SDTM and ADaM have brought great benefits to the entire industry as well as research community, enabling easier data sharing and reuse. In addition, with such standards widely adopted across the industry and a strong trend towards using open-source languages to build solutions for clinical trial analyses, companies started to join efforts and initiated projects such as {admiral}, which is an open-source R package for ADaM datasets creation. Since kicked off by Roche and GSK, it has attracted companies like Amgen, BMS, J&J, Novartis, Pfizer, and extension packages such as {admiralonco}, {admiralophtha}, {admiralvaccine} have been created to meet indication-specific needs. Similarly, R package {roak} originated from Roche, which was designed for SDTM creation, has recently announced a collaboration with CDISC, to build an industry open-source solution for SDTM mapping.

On the contrary, there has been little progress made in adopting the same standard for clinical study reporting, specifically on tables, listings, and graphs (TLGs). Even though a significant portion of clinical trial results are routinely generated (e.g. demographics, safety summary, lab results, etc.) and statistical calculations and methodologies for such results are commonly agreed across the industry (e.g. frequency count, log-rank test in Kaplan-Meier analysis), almost all the pharmaceutical companies have company-specific layout for such TLGs, and implement these TLGs using in-house solutions.

{falcon} is an open-source R package, aiming to build a comprehensive catalog of harmonized tables, listings, and graphs (TLGs) for clinical study reporting. In this paper, we are going to briefly discuss the motivation of having such an industry collaborative effort, its current progress, as well as how we envision this effort for the near- and long-term future.

MOTIVATION

In August 2022, Center for Drug Evaluation and Research (CDER) from US Food & Drug Administration (FDA) proposed an integrated guide for standard safety tables and figures for clinical studies, where a list of suggested safety outputs was provided with implementation considerations. With such guide potentially becoming an industry-wide standard for safety analysis in clinical reporting, pharmaceutical companies were highly engaged in providing comments and suggestions to the guide while it was open for feedback.

The proposed guidance emerged at around the same time when a number of pharmaceutical companies were looking for opportunities to collaboratively develop open-source R packages for clinical reporting. {pharmaverse} was quickly established since its inception as one of the leading effort for industry collaboration on open-source tools, and collaboration on {admiral} was underway with a handful of companies onboard.

Both the proposed integrated guide from FDA and the increasing interest of developing open-source tools together across companies lead to the emergence of an unprecedented opportunity for pharmaceutical companies to collaborate and build solutions for creating TLG outputs. In October 2022, Boehringer Ingelheim, Moderna, Roche, and Sanofi joined efforts together to create an R package called {falcon} under {pharmaverse}, with the aspiration of open-sourcing a catalog of harmonized TLGs for clinical study reporting and simplifying the process of output review, re-use, and meta-analyses.

CURRENT PROGRESS

TEAM SETUP

While kicking off this cross-company collaborative effort, participating companies agreed to adopt existing open-source R packages to build tables and figures from the FDA's integrated guide. Specifically, {rtables} and {tern} packages under the NEST initiative from Roche were proposed to be the main toolkit to be used for {falcon}. A product owner team was formed by one representative from each company, and 12 developers (with one as the team lead) from these companies joined effort to co-develop the {falcon} package.

To effectively manage objectives and priorities, the team adopts agile framework for project coordination. Product owners meet periodically to prioritize features, refine requirements and set and adjust project roadmap, while developers meet weekly in a stand-up meeting to discuss technical issues, progress made in the past week and planned work ahead. GitHub project board and slack channel were established to track development progress as well as to enable effective team communications.

Goals and development priorities set up by product owners were broken down into smaller actionable issues on GitHub by the development team lead, and developers are expected to work on open issues on a self-assignment basis.

DEVELOPMENT PROGRESS

{falcon} website (https://pharmaverse.github.io/falcon/) was launched in June 2023 as a central place to share the development progress, template code, training and workshop materials, as well as answers to frequently asked questions.

As of October 2023, 18 templates were successfully launched in {falcon}. For each template, a ready-to-use code example is provided, with detailed explanation on how each argument in the associated R function could be used. Although majority of the tables are built using NEST packages, specifically {rtables} and {tern}, for some templates, other R-based table engines such as {gtsummary} and {tplyr}, were adopted as well. More templates are being worked on and will be released in the near future.

{falcon} has received increasing awareness from the industry since the kick-off. Representatives from other pharmaceutical companies have reached out and expressed interest to join the effort. Amgen joins the product owner discussions since September 2023 and 4 new developers from different companies have joined the team to co-develop more templates for the package. CDISC also reached out to inquire about future collaborations, with a shared vision on promoting industry standard for clinical reporting.

Most recently in the annual R/Pharma conference, {falcon} was invited to give a hands-on workshop at the {pharmaverse} Friday event. Hundreds of statistical programmers and data scientists participated in the workshop and engaged in the discussion on how to collaboratively make {falcon} better moving forward. The presentation and workshop materials are made freely available on the {falcon} website.

KEY LEARNINGS SO FAR

After one year of this unprecedented cross-company collaboration on creating results for clinical reporting, there are a few learnings we would like to share with the wider community so that it could benefit more people when new collaborative efforts are considered to be initiated.

First, having a well-established industry-wide standard is critical in lowering the collaboration entry point. SDTM and ADaM data models have provided tremendous benefits to the industry and laid out a great basis for collaboration efforts such as {admiral} and {roak}. While FDA's safety integrated guide is not yet a binding requirement for clinical reporting when it comes to safety analyses, having such guidance from the regulatory agency creates a great opportunity for pharmaceutical companies to collaborate and tackle the problem together. Instead of comparing company-specific layouts and trying to convince each company's internal standard implementation department, collaborators focus on one layout proposed by the authority and build solutions towards that layout.

Second, developers are highly motivated to work on open-source project, and it opens new career opportunities for them as well. Making contributions to an open-source project on GitHub is highly valued among the developer's community and building solutions for the entire pharmaceutical industry rather than just one company is also much more attractive to developers and statistical programmers.

Last but not least, building open-source solutions together across pharmaceutical companies is less resource intensive and much more efficient. Instead of having one company investing a team of developers to build code templates for just one company, {falcon} is crowd-sourced by multiple companies and the code templates are shared by all contributors. Managing roadmaps and priorities transparently on GitHub also reduces a lot of frictions and inefficiencies in the process of building solutions.

FUTURE OUTLOOK

As of today, {falcon} is available for download and installation directly from GitHub. We expect it to be made available on CRAN (https://cran.r-project.org/) in the coming months. While we continue developing the {falcon} package, we also expect to collaborate with CDISC closely so that {falcon} aligns with CDISC's Analysis Result Data and Metadata (ARD & ARM) efforts.

As a next step, participating companies plan to start discussing the possibility to expand the scope of work to common analyses in clinical trials and further promoting a harmonized TLG standard for the industry. Moving forward, we hope to engage with more pharmaceutical companies and health authorities in this collaborative effort, with a vision of setting up an industry harmonized TLG standard for clinical reporting, which would replace all internal standards and the implementation is freely accessible for all.

RECOMMENDED LINKS

{falcon} website: https://pharmaverse.github.io/falcon/
{falcon} GitHub repository: https://github.com/pharmaverse/falcon

CONTACT INFORMATION

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

Vincent Shen
Roche
Email: vincent.shen.vs1@roche.com

Jürgen Böhl
Boehringer Ingelheim
Email: juergen.boehl@boehringer-ingelheim.com