Paper OS09

Developing Validated R Packages for Standardized and Efficient Statistical Report Generation at Roche Diagnostics

Dorothee Childs, HMS Analytical Software GmbH, Heidelberg, Germany
Maximilian Kreienbaum, HMS Analytical Software GmbH, Heidelberg, Germany
Christoph Ehret, Roche Diagnostics GmbH, Penzberg, Germany

Introduction

The Value of Diagnostics

The need for better diagnostic solutions to prevent, diagnose and treat diseases is enormous, worldwide:

- 2/3 of diseases are either still not treated adequately or not treated at all;
- 71% of all deaths globally are caused by non-communicable diseases (NCDs), and therefore the medical need is urgent;
- 1/2 of the world lacks access to essential health services;
- only 4% of patients are in clinical trials; we are missing valuable insights and data from 96%.

Diagnostics account for ~70% of clinical decision making, but they are causing only about 2% of total healthcare spending. Hence, it is crucial to develop better diagnostic solutions, and make them available worldwide.

Roche Diagnostics is committed to not only advancing diagnostic solutions that can support healthcare professionals in making critical decisions for their patients' health, but also improving the way these essential tools are integrated into health systems around the world, at a lower cost, while still achieving the best possible outcomes.

In more than 100 countries, Roche Diagnostics offers the industry's most comprehensive in vitro diagnostics testing portfolio combined with clinical decision support, consultancy, digital diagnostics, disease management, laboratory automation, and software solutions to provide integrated solutions for laboratories, clinicians, and healthcare systems.

After development of a new diagnostic assay, its analytical and clinical performance must be demonstrated in clinical studies against pre-specified acceptance criteria, as well as its safety – similar to drug trials.

Analytical performance comprises aspects as e.g. linearity, precision, limit of blank / detection / quantitation that ensure that a lab result is valid and reliable. For the doctor, only the result matters, it must be the same across different manufacturers of an assay.

For clinical performance, the intended use of the assay is specified and demonstrated in the respective target population, against a gold-standard or a predicate device. An example is the intended use of the Elecsys® HCV (Hepatitis C) Duo assay:

"Elecsys® HCV Duo is an immunoassay for the in vitro qualitative detection of hepatitis C virus (HCV) core antigen (HCV Ag) and antibodies to HCV (anti-HCV) in human serum and plasma. The test, in conjunction with other laboratory results and clinical information, may be used to aid in the diagnosis of and the screening for HCV …"

For analytical performance studies, a variety of tools is already available at Roche Diagnostics to analyze most of these highly standardized experiments. Clinical performance studies, though, offer a greater variety of analysis types and their reporting across assays. Hence, there is currently a lack of standardization, causing time-consuming manual analyses and validation procedures.

The Tables, Listings & Graphics catalog TLGcat

Our vision is to provide fast and high-quality reporting for clinical performance studies, embedded in a highly efficient submission process, accelerating time-to-market and hence patients' access to Roche products.

Accordingly, our mission is to:

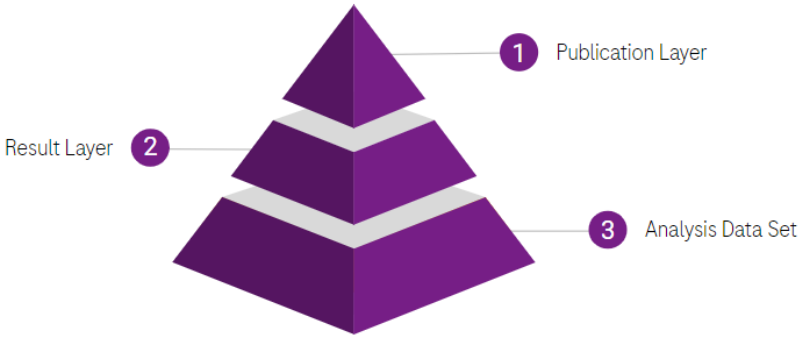
- develop a standardized Tables, Listings & Graphics catalog (TLGcat),
- reconcile it internally with key functions like Clinical Operations and Regulatory Affairs,
- and to automate report generation.

As a guiding principle for the development of the TLGcat R package, we use the following model:

Publication Layer represents the output, i.e. the TLG, as aligned with stakeholders from our disease areas

Result Layer holds every result variable that is needed to fill the Publication Layer's TLG

ADS (analysis datasets) are transformed raw data and follow the specification of a certain TLGcat function



So, how do we get there?

The Roche team for clinical biostatistics has expansive knowledge and experience with statistical designs, analyses and submissions, including the programming skills for tailored analysis. Also, numeric cores are available, usually as R packages developed in house.

However, the team is lacking professional software development skills, especially for:

- Writing code:
    - Running on very different data sets, for different indications
    - Easy to maintain
    - Easy to expand with new features
- Validation of code:
    - Creation of test cases
    - Automatized testing
- Documentation:
    - Fulfilling internal Roche Requirements and bullet proof for external audits

A joint venture between Roche Diagnostics and a highly specialized and experienced software developer like HMS was the obvious choice to bring both strengths and skill sets together.
To this purpose, we developed a validated R package called TLGcat (Tables, Listings, Graphics catalog). The package offers a standardized set of functions to compute frequently requested statistical operations and to present the results in a publication-ready format.

R packages are essential tools in the R programming ecosystem, providing a structured way to bundle together collections of R functions, advanced statistical methods, and other resources. These packages are designed to enhance the functionality of R, making it a powerful tool for data analysis and statistical computing.

At their core, R packages are collections of R functions that can be easily shared and reused. This modular approach allows for efficient code management and reuse, promoting best practices in software development. Many R packages include advanced statistical techniques, enabling users to perform complex data analyzes with ease. This makes R a preferred choice for statisticians and data scientists. Each R package comes with detailed documentation, which includes descriptions of the functions, their parameters, and examples of how to use them.
To ensure reliability and accuracy, R packages often include unit tests which help verify that the functions within the package work as intended, reducing the risk of errors in data analysis.
Some R packages even provide executable analysis templates, which are pre-configured scripts that can be used to perform specific types of analyses and can help save both time and effort, especially for repetitive tasks.

Benefits of Using R Packages

R packages come with everything needed for effective data analysis, including functions, documentation, and tests. This all-in-one format simplifies the process of integrating new tools into existing workflows.

The standardized structure of R packages also makes them easy to integrate into various data analysis processes which is particularly beneficial in agile software development environments, where quick adaptation and iteration are key.

The R community has developed a vast array of packages for documentation, testing, and various analytical tasks. This extensive library of community-contributed packages provides users with a wealth of resources to draw upon, enhancing the capabilities of their analyses.

R packages are well-suited for software engineering (SWE) in agile settings. Their modular nature allows for rapid development, testing, and deployment, aligning well with the principles of agile methodologies.

Regulatory Requirements in the Diagnostics and Pharmaceutical Industry

Regulatory requirements in the healthcare industry are crucial for ensuring the safety, efficacy, and quality of their products. However, meeting these requirements can be a time-consuming process. Interestingly, many of the principles involved overlap with good software engineering practices, which can help streamline compliance efforts.

Requirements engineering is essential in the healthcare industry, just as it is in software development. This involves identifying regulatory needs and ensuring that all aspects of the product meet these requirements. Rigorous testing is a cornerstone of both product development and software engineering. In the healthcare industry, this includes clinical trials, stability testing, and quality control tests to ensure that products are safe and effective. Comprehensive documentation is mandatory, including detailed records of all processes, tests, and results. Good documentation practices ensure traceability and accountability, which are critical for regulatory compliance.

The healthcare industry is subject to specific legal documentation guidelines and obligations, including compliance with regulations from bodies such as the FDA, EMA, and other national regulatory agencies. In addition to external regulations, companies often have their own quality management guidelines. These internal standards help maintain consistency and quality across all products and processes.

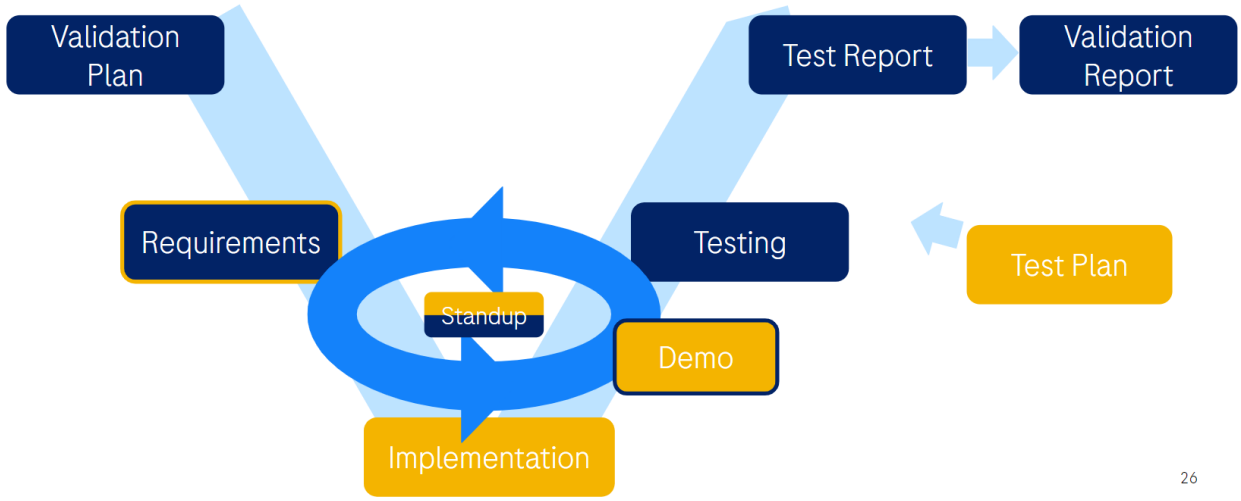
Meeting regulatory requirements demands a substantial investment of time and resources, which can be a major challenge for diagnostics and pharmaceutical companies. However, by adopting good software development practices, such as agile methodologies, continuous integration, and automated testing, these companies can optimize their processes. These practices help streamline documentation, improve testing efficiency, and ensure that requirements are consistently met.

In summary, while regulatory requirements in the healthcare industry are essential and time-consuming, they can be effectively managed by leveraging good software engineering

practices. This approach not only ensures compliance but also enhances the overall efficiency and quality of product development.

In recent years, several community initiatives have emerged in order to support R-based clinical reporting with a focus on standardized tables and figures, as well as compliance with regulatory requirements [1, 2, 3]. Such initiatives provide valuable assets for standardized views that can be integrated into customized analysis reports and be subject to further validation on the report side. In the following, we present our experiences in developing a customized tables, listings and graphics catalog for in-house expert users in a highly regulated environment.

Methods



As usual for validated software our approach to requirements, implementation, and testing adhered to the V-model, in this case with an agile component, ensuring a comprehensive and methodical, yet flexible, development process.

For the requirements, the focus was on fostering a highly collaborative and interactive environment. This phase saw close alignment between HMS developers and Roche experts. Success here hinged on having committed contacts who not only had their time blocked but were also able to utilize it effectively. These key players maintained close contact with both developers and internal experts, ensuring seamless communication. The process was supported by closely timed feedback cycles, ensuring continuous alignment with Quality Management (QM). Participation in sprint-planning and stand-up meetings further solidified this alignment, making the processes a key success factor.

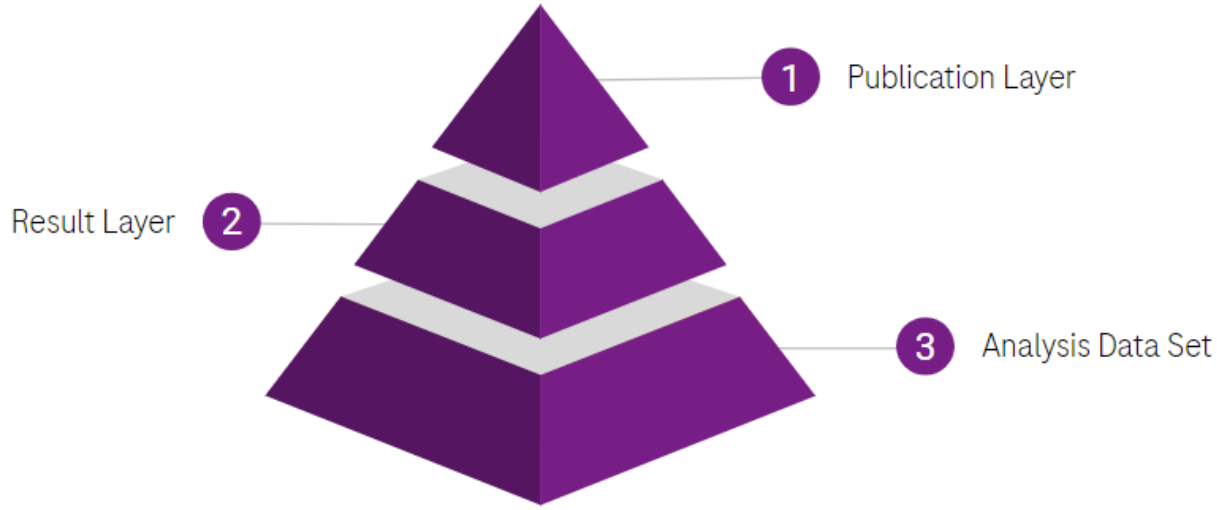
During Implementation, trust in people emerged as the pivotal success factor. Developers were trusted to deliver their best work within the available time, while expert users were relied upon to maintain realistic expectations. Quality assurance was reinforced by unit testing alongside the main implementation, complementing the validation tests. Regular code demos allowed developers to showcase progress, facilitating quick feedback from expert users who used very good test data allowing for early catching of bugs and

inconsistencies. Realistic availability of experts for testing and the flexibility to prioritize and descope when necessary were critical to maintaining momentum and focus.

In the Testing phase, processes once again played a central role. Close collaboration with QM ensured that the correct documents, complete with the right attachments, were filled out accurately and submitted on time. A clearly defined scope helped limit the validation efforts, allowing for a more efficient process. The close coupling between development and validation teams ensured seamless integration and swift resolution of any issues. Additionally, the use of high-quality test data that accurately represented actual study data was crucial. Tools such as automatic document rendering within the system and largely automated tests, utilizing standard unit test frameworks like testthat, significantly increased reproducibility. The Roche-approved internal package repository provided a secure and efficient way to manage all software packages, again limiting the scope of the effort. Validation testing was mainly conducted by automated tests implemented using the R package testthat. Reports showing the number of successful tests were produced using the R package *rmarkdown*.
Ultimately, the success of the project was built on strong processes, trust in people, continuous alignment, and effective use of advanced tools.

Results

The TLGcat package consists of a three layer structure.

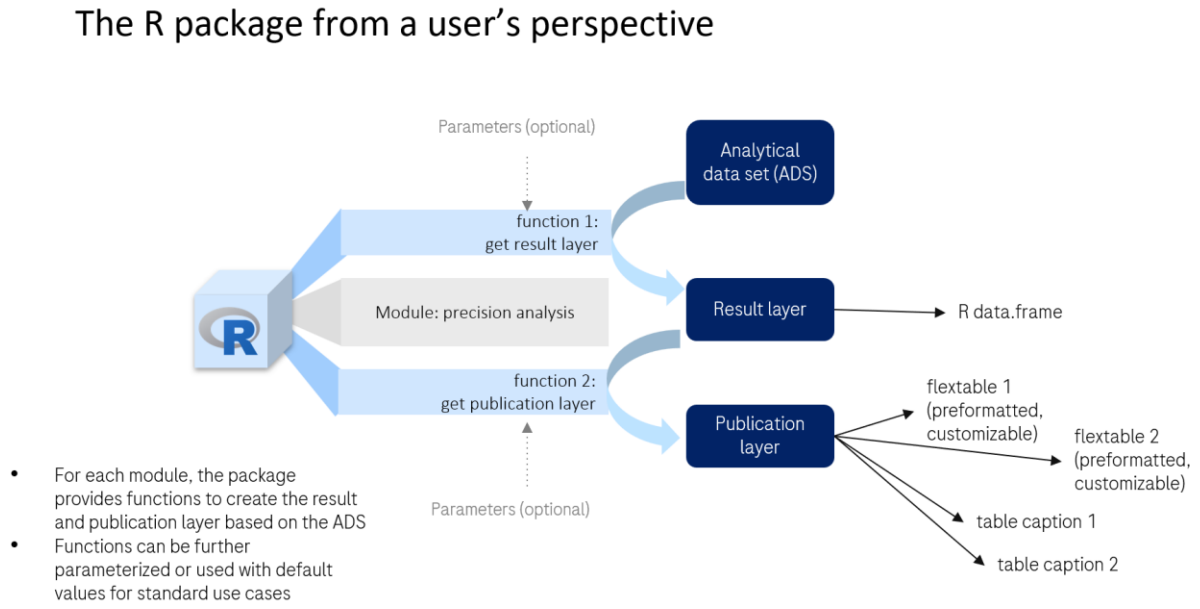


The publication Layer represents the output, i.e. the TLG, as aligned with stakeholders from different disease areas. The Result Layer holds every result variable that is needed to fill the Publication Layer's TLG.The ADS (analysis datasets) layer consists of transformed raw data and follows the specification of a certain TLGcat function.

As of today, the R package consists of modules for Reproducibility, Reference Ranges, and Qualitative Method Comparisons. Tables produced by the package follow a

style guide that defined publication ready requirements for fonts, margins, headers etc., considering also section 508 of the US Dept. of Justice.

The R package provides individual functions for each layer and statistical module which can be further parameterized if needed by the users.



Example for precision:

Table X: Summary table of reproducibility using *REML/ANOVA*, *including/excluding* outliers.

Sample	Mean	N	*N (outliers removed)*	Repeatability SD	Repeatability CV [%]	Between-Run SD	Between-Run CV [%]	Between-Day SD	Between-Day CV [%]	*Within-Site Within-Lot SD*	*Within-Si Within-Lo CV [%*

*negative variance estimates were set to 0.

Table X: Variances table for reproducibility using *REML/ANOVA, including/excluding* outliers.

Sample	Mean	N	*N (outliers removed)*	DF	SD Estimate	SD 95% LCL	SD 95% UCL	CV Estimate	CV 95% LCL	CV 95% UCL	Claim Fullfilled yes/no

*negative variance estimates were set to 0.

Example for reference ranges:

Table X: Quantile estimation for *<assay> [<units>]* with *<90/95%>* confidence interval using *<method, e.g. non-parametric, symmetric>*

grpvar1 (disease grp)	*grpvar2 (sex)*	*grpvar3 (agegrp)*	N	Quantile	Estimate	LCL	UCL
A	female	young	150	*Minimum*		n.a.	n.a.
A	female	young	150	*0.5%*			
A	female	young	150	*1%*			
A	female	young	150	*2.5%*			
A	female	young	150	*5%*	52.1	*	*
A	female	young	150	*10%*	56.7	40.2	12.6
A	female	young	150	*25%*			
A	female	young	150	*Median*	70.5	48	14.3
A	female	young	150	*75%*			
A	female	young	150	*90%*	54.8	64.4	17.2
A	female	young	150	*95%*	57.2	*	*
A	female	young	150	*97.5%*			
A	female	young	150	*99%*			
A	female	young	150	*99.5%*			
A	female	young	150	*Maximum*			
A	female	old	905	*Minimum*			
A	female	old	905	*2.5%*			

Conclusion

First experiences at Roche Diagnostics show faster programming at even higher quality, with less efforts needed for validation. For a standard precision analysis for example, traditionally ~1500 lines of code were written, collapsing with the help of the TLGcat now to only 50 lines that are also less error prone. For our future use cases, we expect a similar gain in efficiency and quality. Next steps are to collect experience from future submissions with the TLGcat to see if regulators' needs can be covered, and afterwards it is planned to make the TLGcat publicly available.

Efficiency in our development and validation approach hinged on several key factors. First and foremost, close coordination on requirements was vital, ensuring that everyone involved was on the same page at all times. Adopting a working mode that incorporates an agile framework can significantly enhance this process. Tailoring these agile processes to fit the team's specific needs and then living them together cultivates an approach that is both satisfying and productive.

Additionally, maintaining a shared backlog was crucial. This backlog served as a centralized repository to collect next steps in all directions, benefitting both the experts and the development team. By harmonizing efforts through these practices, we were able to navigate the complexities of development and validation with greater ease and success.

Overall, the collaboration of the clinical biostatistics team at Roche Diagnostics with the HMS team turned out to be very successful and we are looking forward to the next steps in this joint endeavor.

References

[1] R package falcon and cardinal: cardinal (pharmaverse.github.io)

[2] pharmaverse
[3] R package valtools: GitHub - phuse-org/valtools: Validation framework for R packages used in clinical research and drug development.