Paper OS10

# {admiralmetabolic}: A collaborative Journey to Develop an Admiral Extension Package for Metabolic Data

Anders Askeland, Novo Nordisk, Aalborg, Denmark
Olga Starostecka, Boehringer Ingelheim, Germany

## ABSTRACT

In recent years, the field of pharmaceutical data science has witnessed a significant shift towards the R programming language for regulatory analysis and reporting[1–3] . While initially focused on outputs, comprehensive tooling now enables R utilization across the entire clinical trial data workflow, including SDTM transformation using {sdtm.oak}[4], ADaM dataset creation using {admiral}[5], results reporting using tools such as {rtables}[6] and {gtsummary}[7], and dynamic data visualization facilitated by packages like {teal}[8] and {davinci}[9]. Despite the maturity of {admiral} (ADaM In R Asset Library), its actual implementation remains sparse in specialized therapeutic areas like metabolic diseases, gaps in existing tooling and documentation make implementation daunting. To address this, collaborators from Novo Nordisk, Roche, Boehringer Ingelheim, and Novartis developed {admiralmetabolic}, an extension package for metabolic-specific ADaM datasets. Developing {admiralmetabolic} offered a unique opportunity to learn the complexities of the {admiral} ecosystem, with benefits such as accessible starting points, robust community support, continuous knowledge sharing, and cross-industry collaboration. This paper explores the journey of creating this package, detailing its key functionalities, the collaborative model that made it possible, and the future roadmap. We provide practical use cases and share insights on the value of open-source contribution as both a learning strategy and a method for advancing industry-wide standards.

## INTRODUCTION

The pharmaceutical industry is undergoing a data analysis paradigm shift, with R emerging as the new standard for clinical trial reporting and regulatory submissions. This transition is driven by R's open-source nature, which promotes transparency, reproducibility, and access to a vast ecosystem of cutting-edge statistical and data reporting tools. Initiatives like pharmaverse are accelerating this change by fostering a collaborative environment where companies and individuals build and maintain a suite of validated interoperable R packages designed specifically for the clinical workflow.

At the core of this ecosystem is the {admiral} (ADaM In R Asset Library) package. It provides a modular and opinionated framework for creating ADaM datasets compliant with the CDISC standards. However, the diverse and specialized nature of clinical research means that a one-size-fits-all approach is insufficient. Therapeutic areas such as metabolic diseases involve unique endpoints and complex derivations that require tailored tooling and documentation. This gap between the core {admiral} framework and the specific needs of metabolic research was the catalyst for the creation of {admiralmetabolic}, a {admiral} extension package focusing on metabolic diseases. The package has an initial focus on the obesity therapeutic area and is designed to complement {admiral} by offering reusable derivations, templates, and datasets. It is available on CRAN (Apache-2.0 license) with R ≥ 4.1 and a dependency strategy aligned to the released {admiral} core, ensuring reproducibility and stable integration into existing pipelines. This paper details the development of this extension package, presenting it as a case study in effective open-source collaboration and a practical guide for users.

## THE COLLABORATIVE DEVELOPMENT MODEL

The {admiralmetabolic} package is an outcome of the power of cross-company collaboration in an open-source setting. The initiative brought together developers from four otherwise competing pharmaceutical companies such as Novo Nordisk, Roche, Boehringer Ingelheim, and Novartis under a shared goal: to build a high-quality, reusable tool and documentation for metabolic specific ADaM datasets for the entire industry. The effort leveraged pharmaverse community infrastructure and norms, including transparent governance, shared coding standards, and open review practices to maximize reuse across sponsors.

At the core of this development model was a lightweight but effective governance model, which was key to the project's success. Work was coordinated through the GitHub repository, utilizing features like Issues for feature requests and bug reports, Pull Requests for code contributions, and Projects for roadmap planning. Communication flowed through dedicated Slack channels, allowing for real-time problem-solving and knowledge sharing. A crucial aspect of the workflow was the peer-review process where every piece of code was reviewed by at least one other developer from a different company, ensuring high code quality and alignment with the {admiral} core principles. This model fostered a sense of shared ownership and collective responsibility.

## VALUE PROPOSITION

Engaging in this collaborative effort offered a compelling return on investment for both individual developers and their parent companies.

- **For Developers**: It provided an unparalleled learning opportunity. Contributors gained deep, practical expertise in the {admiral} ecosystem, sharpened their R programming skills in a best-practice environment (e.g., package development, unit testing), and built a strong professional network.
- **For Companies**: Contributing to {admiralmetabolic} was a strategic investment. It allowed companies to influence the development of an industry-standard tool, ensuring it met their specific needs. Furthermore, it helped build internal expertise and upskill their programming teams, reducing reliance on proprietary software and fostering a culture of innovation. By sharing the development burden, each company achieved a more robust and feature-rich solution than they could have built alone.

## CHALLENGES AND LESSONS LEARNED

The developmental journey was not without its challenges, each of which provided valuable lessons.

- **Technical Alignment**: One of the primary technical hurdles was ensuring that the extension remained synchronized with the core {admiral} package, which was itself under active development. This required communication with the {admiral} core team and a commitment to adapting the {admiralmetabolic} code to upstream changes. The lesson learned was the importance of building extensions in a modular way that anticipates and isolates potential breaking changes.
- **Standardizing Across Organizations**: While all participating companies follow CDISC standards, subtle differences in interpretation and internal implementation conventions exist. Reaching a consensus on function design, parameter naming, and the scope of derivations required extensive discussion and a willingness to compromise for the sake of creating a universally applicable tool.
- **Resource Allocation**: As contributions were made on top of regular job responsibilities, managing time and maintaining momentum was a constant challenge. The key to overcoming this was breaking down large features into smaller, manageable tasks and celebrating small wins to keep the team motivated and engaged.

## KEY FEATURES AND FUNCTIONALITY

{admiralmetabolic} aims to provide tailored functionalities and documentation beyond the core {admiral} package, including: metabolic specific SDTM test data to facilitate development and validation, detailed vignettes and corresponding template programs, and several purpose-built derivation functions for vital-sign derivation.

In relation to documentation, {admiralmetabolic} is distributed with comprehensive documentation and example data, including:

- **Vignettes**: The package includes long-form tutorials (vignettes) that provide step-by-step instructions for creating common ADaM datasets like ADVS, ADLB, and ADCOEQ. These are not just technical manuals; they are educational resources that explain the context and best practices.
- **Example Data**: The inclusion of metabolic-specific SDTM test data for LB (Laboratory), VS (Vital Signs) and QS (Questionnaires) domains allows users to run the examples directly and serves as a clear specification for the data structures the functions expect.

In relation to functions, the current functions included in {admiralmetabolic} are designed to simplify common, complex derivations. The primary value-add is abstracting away boilerplate code, reducing the chance of error, and improving the readability and maintainability of analysis scripts. Currently, {admiralmetabolic} includes two functions; `derive_param_waisthgt()` to calculate waist height ratios and `derive_param_waisthip()` to calculate waist and hip circumference ratio.

### Example of the `derive_param_waisthgt()` function

Consider the task of calculating the Waist-to-Height. Without {admiralmetabolic}, a programmer would use the `admiral::derive_param_computed()` function, which requires one to manually specify the formula, potential constant units, and takes no consideration of the collected units:

```
derive_param_waisthgt(
    dataset = advs,
    by_vars = exprs(USUBJID, VISIT),
    wstcir_code = "WSTCIR",
    height_code = "HEIGHT",
    set_values_to = exprs(
        PARAMCD = "WAISTHGT",
        PARAM = "Waist to Height Ratio"
    ),
    constant_by_vars = exprs(USUBJID),
    get_unit_expr = admiral::extract_unit(PARAM)
)
```

With {admiralmetabolic}, the dedicated `derive_param_waisthgt()` function simplifies this significantly, where it handles the underlying logic, including unit conversions, internally:

```
derive_param_waisthgt(
  dataset = advs,
  by_vars = exprs(USUBJID, VISIT),
  wstcir_code = "WSTCIR",
  height_code = "HEIGHT",
  set_values_to = exprs(
    PARAMCD = "WAISTHGT",
    PARAM = "Waist to Height Ratio"
  ),
  constant_by_vars = exprs(USUBJID),
  get_unit_expr = admiral::extract_unit(PARAM)
)
```

This simplification is a recurring theme, allowing programmers to focus on the "what" rather than the "how".

### Example of the ADLB vignette

In addition to more concrete workflows and functionalities, {admiralmetabolic} also provides documentation of practical use cases, where it includes vignettes of how to calculate common derivations used in metabolic trials. To demonstrate how the package functions in a real-world workflow, the ADLB vignette outlines deriving the HOMA-IR and Fatty Liver Index (FLI), which require data from multiple sources.

First, to derive HOMA-IR:

```
adlb <- adlb %>%
  derive_param_computed(
    by_vars = exprs(USUBJID, AVISIT, AVISITN, ADT, ADY),
    parameters = c("INSULIN", "GLUC"),
    set_values_to = exprs(
      AVAL = AVAL.INSULIN * AVAL.GLUC / 22.5,
      PARAMCD = "HOMAIR",
      PARAM = "Homeostasis Model Assessment - Insulin Resistance"
    )
  )
```

To derive FLI, we first need to add BMI and waist circumference (WSTCIR) from ADVS.

```
adlb <- adlb %>%
  derive_vars_transposed(
    advs,
    by_vars = exprs(USUBJID, ADT),
    key_var = PARAMCD,
    value_var = AVAL,
    filter = PARAMCD %in% c("BMI", "WSTCIR")
  )
```

Following this, we can derive the FLI score

```
adlb <- adlb %>%
  derive_param_computed(
    by_vars = exprs(USUBJID, AVISIT, AVISITN, ADT, ADY, BMI, WSTCIR),
    parameters = c("TRIG", "GGT"),
    set_values_to = exprs(
      AVAL = {
        lambda <- 0.953 * log(AVAL.TRIG) + 0.139 * BMI + 0.718 * log(AVAL.GGT)
                  + 0.053   * WSTCIR - 15.745
        (exp(lambda) / (1 + exp(lambda))) * 100
      },
      PARAMCD = "FLI",
      PARAM = "Fatty Liver Index"
    )
  )
```

This example demonstrates how {admiralmetabolic} not only provides convenience functions, but also established patterns and examples for performing multi-step metabolic derivations within the {admiral} framework.

## FUTURE ROADMAP AND COMMUNITY INVOLVEMENT

The development of {admiralmetabolic} is an ongoing, community-driven effort. The future roadmap is transparently managed on GitHub and reflects the evolving needs of metabolic specific clinical trials.

Key features planned for future releases include:

- **Oral Glucose Tolerance Test (OGTT) Support**: This is a high-priority feature that will introduce functions to calculate key OGTT parameters, such as Area Under the Curve (AUC) for glucose and insulin, which are fundamental endpoints in diabetes research.
- **Glycemic Status Derivation**: A function is planned to classify subjects into glycemic categories (e.g., normal, pre-diabetes, diabetes) based on laboratory values like HbA1c and fasting glucose, according to guidelines from organizations like the American Diabetes Association (ADA).
- **Cardiovascular Risk Scores**: To broaden the package's applicability, functions to calculate established cardiovascular risk scores (e.g., Framingham Risk Score) are under consideration.
- **Weight Loss Plateau Derivations:** Introduce a method to derive time-to-plateau (e.g., ≤5% weight change over a contiguous intervention window), delivered via a vignette and potential helper functions.

### A CALL TO ACTION

The success of {admiralmetabolic} and the wider pharmaverse ecosystem depends on community participation. We encourage users and developers to get involved. Whether it is by reporting a bug, requesting a new feature, contributing to documentation, or submitting code, every contribution is valuable. By working together, we can build a comprehensive, open-source toolset that accelerates the analysis of metabolic clinical trial data.

## CONCLUSION

The {admiralmetabolic} package is more than a collection of R functions and vignettes; it is a case study in open-source collaboration within the pharmaceutical industry. It demonstrates how sharing expertise and development effort can lead to the creation of high-quality, standardized tools that benefit everyone. The journey of its creation has provided an effective, hands-on learning environment for its contributors, deepening their knowledge of R, {admiral}, and ADaM standards. As the package continues to evolve with community input, we hope that it will become a resource for anyone working with metabolic clinical trial data in R.

## REFERENCES


1. R Submissions Working Group. https://rconsortium.github.io/submissions-wg/.
2. pharmaverse. https://pharmaverse.org/.
3. Leng, L., Patel, H. & Jingyuan, C. Roche's end-to-end R submission for a new drug application. *Roche's end-to-end R submission for a new drug application* https://posit.co/blog/roches-first-end-to-end-r-journey-to-submission/ (2024).
4. Knoph, A. S., Larsen, S. F. & Bilgrau, A. E. Novo Nordisk: Journey to an R-based FDA submission. *Novo Nordisk: Journey to an R-based FDA submission* https://posit.co/blog/novo-nordisks-journey-to-an-r-based-fda-submission/ (2023).
5. Ganapathy, R. *et al.* sdtm.oak: SDTM Data Transformation Engine. (2024).
6. Straub, B. *et al.* admiral: ADaM in R Asset Library. (2025).
7. Becker, G. *et al.* rtables: Reporting Tables. (2025).

8. Sjoberg, D. D. *et al.* gtsummary: Presentation-Ready Data Summary and Analytic Result Tables. (2025).
9. Kaledkowski, D. *et al.* teal: Exploratory Web Apps for Analyzing Clinical Trials Data. (2025).
10. Glauß, I. & Döring, S. davinci: Dynamic Visualization for Clinical Insights. (2025).
11. Askeland, A. *et al.* admiralmetabolic: Metabolism Extension Package for ADaM in 'R' Asset Library. (2025).

## ACKNOWLEDGMENTS

We would like to express our sincere gratitude to the entire {admiralmetabolic} development team for their dedication, hard work, and collaborative spirit that made this project possible. The main development team for the first releases (v. 0.1 and v. 0.2) included: Anders Askeland (Novo Nordisk), Andrii Yurovskyi (Roche), Kathrin Flunkert (Roche), Edoardo Mancini (Roche), Shunsuke Goto (Novartis), Siddhesh Pujari (Novo Nordisk), Sonali Das (Novo Nordisk). Olga Starostecka (Boehringer Ingelheim), Vang Le-Quy (Novo Nordisk), and Keita Takahashi (Roche). We would also like to thank the core development team of {admiral} for their support and collaboration throughout the development of this package. Lastly, we extend our gratitude to our employers for fostering an environment that encourages and supports contributions to open-source software.

## CONTACT INFORMATION

**Author Name**: Anders Askeland
**Company**: Novo Nordisk
**Email**: IAKD@novonordisk.com
**Website**: https://www.novonordisk.com/

**Co-Author Name**: Olga Starostecka
**Company**: Boehringer Ingelheim
**Email**: olga.starostecka@boehringer-ingelheim.com
**Website**: https://www.boehringer-ingelheim.com/