Paper OS13

Open Source as a Strategic Enabler for EMA Submissions: An R Shiny Approach

Marckenley Mercie, Bristol Myers Squibb, Madison, NJ

ABSTRACT

The rise of open-source technologies is redefining how organizations approach regulatory submissions. The European Medicines Agency (EMA) requires Risk Management Plan (RMP) tables as part of marketing authorization applications. Traditionally, these outputs have been developed using SAS in processes that are time-consuming and resource-intensive, limiting efficiency across assets. To address this, we developed an R Shiny application to modernize the creation of RMP tables while fitting within existing workflows. The tool automates table generation, shortens timelines, improves consistency, and strengthens collaboration between programming, statistical, and regulatory teams. Its use across different assets demonstrates how open-source solutions can scale effectively while delivering measurable value. This paper highlights how an open-source solution transformed the delivery of EMA-required RMP tables. By improving workflows for both development and submission, it illustrates how R Shiny can drive efficiency while advancing open-source adoption in clinical research.

INTRODUCTION

Risk Management Plans (RMPs) play a central role in regulatory submissions to the European Medicines Agency (EMA). They provide a structured framework for documenting the identification, characterization, monitoring, and minimization of a medicinal product's important risks over its life cycle, as outlined in EMA guidance (European Medicines Agency, 2018). This guidance specifies expectations for the presentation of clinical exposure and adverse event summaries, which support regulatory evaluation of a product's benefit–risk profile.

As medicinal products progress through development and post-authorization phases, RMPs are frequently updated to reflect emerging safety data. These updates often require regeneration of supporting tables, increasing the demand for efficient, reproducible, and scalable reporting workflows. Traditional approaches can become resource-intensive when applied repeatedly across multiple assets and submission cycles. To address these challenges, we explored the use of open-source technologies to modernize RMP table generation while remaining aligned with established regulatory processes. Leveraging R Shiny, we developed a web-based application designed to streamline the creation of RMP tables, improve consistency across submissions, and support more efficient regulatory workflows. The following section provides an overview of the EMA requirements that inform the structure and content of RMP safety and exposure outputs.

RMP TABLE REQUIREMENTS

The European Medicines Agency (EMA) provides clear guidance on the format and content of Risk Management Plan (RMP) tables, with particular emphasis on clinical exposure and safety outputs. For clinical exposure, the EMA recommends that data be summarized at an aggregated level across all relevant clinical trials, rather than on a study-by-study basis unless otherwise justified (European Medicines Agency, 2018). These summaries are intended to characterize the extent and distribution of patient exposure underlying the safety database and to support the interpretation of adverse event findings.

Cumulative for all indications (person time)

Duration of exposure	Patients	Person time
e.g. <1 m		
1 to <3 m		
3 to <6 m		
≥6 m etc.		
Total person time		

<Indication>

Duration of exposure	Patients	Person time
e.g. <1 m		
1 to <3 m		
3 to <6 m		
≥6 m etc.		
Total person time for indication		

Figure 1: Duration of Exposure

Figure 1 presents an example of a clinical exposure table adapted from the EMA guidance. The table summarizes cumulative exposure across all indications, as well as stratified by indication, using predefined duration intervals. Key elements include counts of exposed subjects and corresponding person-time, which together provide regulators with essential context on both the magnitude and duration of exposure across the development program. In addition to overall exposure, the EMA guidance notes that exposure summaries by relevant subgroups—such as age group and gender—may be required to support interpretation of safety findings. These subgroup summaries follow the same structural principles as the overall exposure tables and extend the format to identify populations with limited exposure or potential differential risk profiles.

While the EMA guidance does not prescribe specific statistical metrics for safety summaries, it emphasizes clear and concise presentation of adverse event data to support regulatory assessment. Tabulated safety outputs are encouraged where they improve interpretability, particularly when summarizing adverse events by severity, seriousness, or other clinically meaningful dimensions.

In practice, companies may supplement standard safety summaries with additional metrics that enhance interpretation of risk in the context of clinical exposure. For example, exposure-adjusted incidence rates (EAIRs) and comparative risk measures may be used to contextualize adverse event frequency across treatment groups or development stages. When presented alongside cumulative exposure summaries, such metrics can support a more robust characterization of safety patterns while remaining aligned with the overall intent of the RMP guidance.

WHY OPEN SOURCE?

The preparation of RMP tables requires adherence to regulatory guidance while supporting repeated updates across a medicinal product's lifecycle. As described in the preceding sections, EMA guidance emphasizes structured, consistent presentation of clinical exposure and safety outputs to support regulatory assessment. In practice, however, organizations often support multiple assets simultaneously, each with its own development timelines and submission milestones. The generation of RMP tables represents a substantial portion of the overall submission timeline, particularly when similar outputs must be reproduced across assets and submission cycles. This operational burden highlights the need for approaches that can accelerate table generation while maintaining consistency and regulatory alignment.

An application-based approach was selected to address these challenges by standardizing RMP table layouts and generation logic across submissions. Standardizing table layouts allows an application to enforce consistent structure, formatting, and calculation rules, reducing variability that may arise when similar tables are developed independently across studies or programs.

R Shiny was chosen as the implementation platform due to its ability to create a user-friendly interface. Shiny applications enable complex data wrangling to be implemented in the back-end, allowing users to generate RMP tables without direct interactions with the underlying code. This approach lowers the barrier to adoption for users with limited experience in R while preserving flexibility for advanced customization when needed. By integrating R Shiny into existing programming and validation workflows, organizations can benefit from the transparency, extensibility, and community support of open-source tools while remaining compliant with regulatory expectations. Effectively, R Shiny serves not only as a technical solution, but as a strategic enabler for standardized, efficient, and reproducible RMP table generation.

RMP APPLICATION OVERVIEW

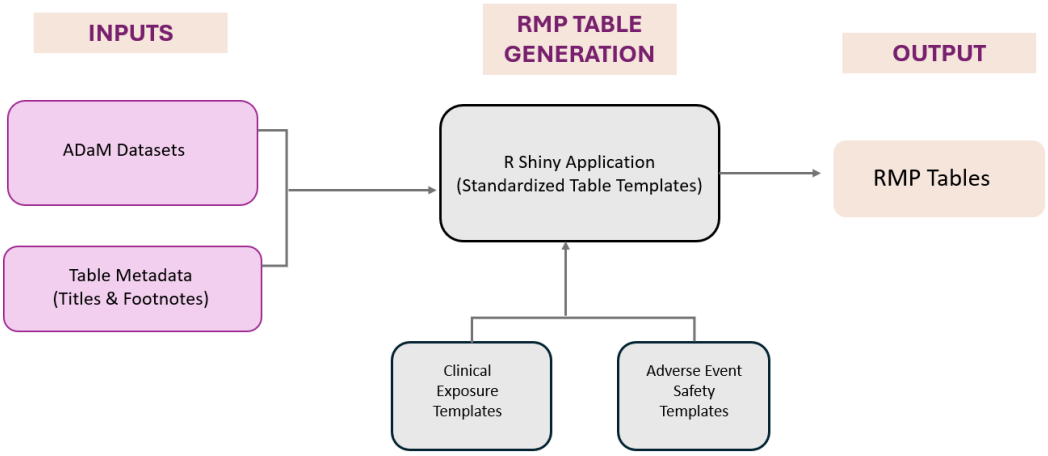

Figure 2: RMP Table Workflow

Figure 2 illustrates the high-level workflow used to generate Risk Management Plan (RMP) tables through the R Shiny application. The application was designed to centralize RMP table generation within a standardized workflow, enabling reliable production of EMA-aligned outputs across multiple assets and submission cycles.

CDISC-compliant ADaM datasets serve as the primary data input to the application. Leveraging ADaM datasets ensures that key variables and parameters required for RMP tables are already derived and standardized prior to table generation. As a result, the application minimizes the need for additional table-specific derivations within the reporting workflow, thereby reducing complexity in the back-end logic. In addition to analysis data, the application incorporates controlled table metadata, including titles and footnotes, which are maintained separately from the datasets. This separation allows standardized table layouts to be reused while supporting asset-specific content and regulatory context.

These inputs feed into the core R Shiny application, which applies predefined table templates through server-side R programs. The server-side logic consists of modular R programs that implement standardized templates for clinical exposure and adverse event safety tables. These programs encapsulate the methods required to calculate summary statistics and apply consistent formatting in accordance with regulatory expectations.

The final output of the workflow consists of RMP tables that are ready for regulatory use and consistent in structure across submissions. By centralizing table logic and presentation within the application, this approach reduces variability introduced by independent table development and supports efficient regeneration of outputs as data are updated over time.

VALIDATION STRATEGY

Validation of RMP tables generated by the R Shiny application follows a traditional double-programming approach consistent with established reporting practices. The Shiny application integrates with existing validation processes, allowing outputs from open-source tools to be verified using established methods.

To support this approach, the application produces XPT files corresponding to the generated RMP tables. These XPT datasets serve as a standardized, system-independent representation of the table outputs and provide a direct input for downstream validation activities. Validation is performed using SAS programs that independently recreate the required summaries based on the same ADaM datasets and table specifications. Discrepancies identified during validation are investigated and resolved prior to finalization of regulatory outputs.

CONSIDERATIONS AND FUTURE ENHANCEMENTS

Execution and deployment of the R Shiny application within a regulated environment require several key considerations. Operation within a GxP-compliant R environment is necessary to ensure appropriate controls around system access, execution, and change management. Establishing and maintaining such an environment is critical to supporting regulatory confidence in outputs generated using open-source technologies. In addition, audit trail processes are essential for traceability and reproducibility. The use of validated R packages further ensures that analytical dependencies are well understood, documented, and suitable for use in regulated reporting workflows. These foundational governance elements are summarized in Figure 3 as the current operational foundations supporting the application.

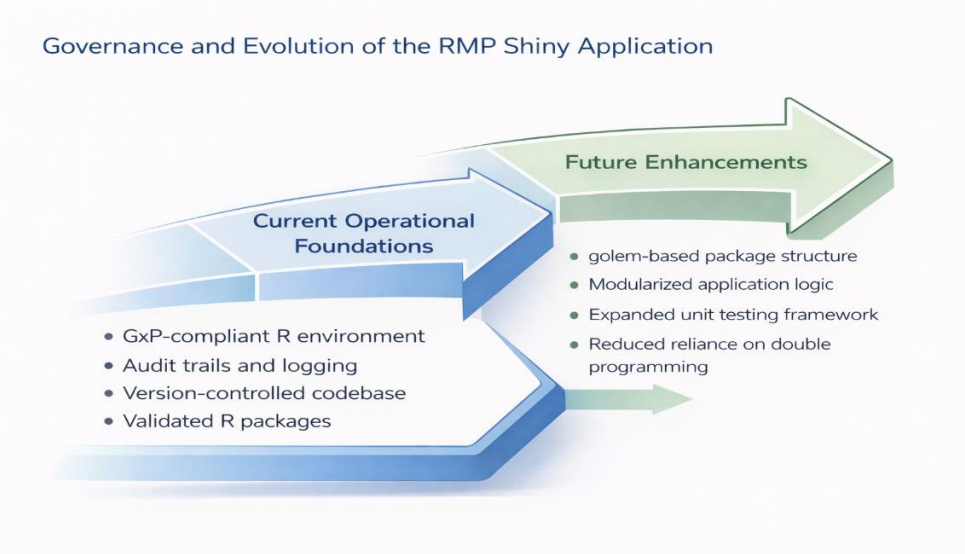

Figure 3: Governance and Evolution of RMP Shiny Application

Future enhancements to the application include formalizing the codebase into a structured R package using a *golem* framework. This approach would further modularize application components, improve maintainability, and support clearer separation between interface logic and analytical functions. In parallel, expanded unit testing of the package structure represents an opportunity to strengthen confidence in table generation logic. Over time, a robust testing framework may reduce reliance on traditional double programming for certain outputs, potentially accelerating workflows while maintaining appropriate quality controls. As illustrated in Figure 3, these enhancements build directly upon the existing governance framework, reflecting an evolutionary path rather than a replacement of established controls.

BENEFITS AND ADOPTION

Implementation of the R Shiny application reduced development effort associated with RMP table generation by eliminating the need to remediate existing programs or develop asset-specific table logic from scratch. By centralizing standardized table layouts and business rules within a single workflow, the application minimizes repetitive programming activities and supports consistent production of regulatory deliverables across assets. Thus, enabling rapid regeneration of RMP tables as data evolves, allowing teams to focus effort on validation activities and cross-functional review rather than table construction.

The application's user interface abstracts underlying implementation details, lowering the barrier to use for individuals who do not routinely program in R. Users can generate RMP tables through a controlled interface without direct interaction with the underlying code, supporting broader usability while maintaining standardized outputs. Collectively, these benefits improve consistency, reduce variability across submissions, and enhance efficiency within regulatory reporting workflows.

The application has been adopted across multiple assets, demonstrating scalability and reuse within the organization. Importantly, adoption of the open-source solution integrates seamlessly with existing SAS-based validation workflows. Continued use across assets provides ongoing feedback that informs incremental improvements to standardized templates and workflows, supporting sustainable adoption and evolution of the solution over time.

CONCLUSION

Risk Management Plan (RMP) tables are a critical component of EMA submissions and require consistent, structured presentation of clinical exposure and safety information throughout a product's lifecycle. As regulatory programs expand across multiple assets and submission cycles, the ability to generate these outputs efficiently and reproducibly becomes increasingly important.

This paper has described the development and implementation of an R Shiny application that supports standardized generation of EMA-aligned RMP tables using CDISC-compliant ADaM datasets and controlled table metadata. By centralizing table logic, layouts, and governance within a single workflow, the application reduces variability across submissions while integrating seamlessly with established validation practices.

Adoption of this open-source solution across multiple assets demonstrates that modern technologies can be applied effectively within regulated reporting environments when supported by appropriate controls and validation strategies. Collectively, this work illustrates how R Shiny can serve as a scalable and sustainable approach for standardizing regulatory outputs while advancing open-source adoption in clinical research.

REFERENCES

European Medicines Agency. (2018, October 31). *Guidance on the format of the Risk Management Plan (RMP) in the EU – integrated format.* Retrieved from European Medicines Agency: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guidance-format-risk-management-plan-rmp-eu-integrated-format-rev-201_en.pdf

Mercie, M., & Tai, D. (2023). Risk Management Plan (RMP) Tables using R. *PHUSE US 2023 Conference Proceedings* (p. PAP_PP03). United States: Pharmaceutical Users Software Exchange (PHUSE).

ACKNOWLEDGMENTS

Portions of this manuscript were refined using AI-assisted tools to support language editing and clarity. All ideas, analyses, technical content, interpretations, and conclusions originate from the author, who retains full responsibility for the content presented herein.

Many thanks to the Hematology, Oncology, and Cell Therapy organization at Bristol Myers Squibb for their continued support.

CONTACT INFORMATION

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

Marckenley Mercie
Bristol Myers Squibb
Email: marckenley.mercie@bms.com