Paper AD01

H2H Portal - A Shiny Automation of Reproducible and Traceable Analysis and Reporting Using Metadata Framework

Julian Schmocker, MSD, Zurich, Switzerland
Qian Wang, MSD, Brussels, Belgium
X. Gregory Chen, MSD, Zurich, Switzerland

ABSTRACT

The process to generate reproducible and traceable statistical outputs is complex. Typically, it requires inputs from cross-functional stakeholders, captured in various documents, and implemented potentially with different programming software, all within a tight timeline.

This paper shares the journey and lessons of developing an R Shiny app, H2H Portal, an internal end-to-end automation to provide additional clinical statistics beyond the clinical study reports in support of economic evaluation for Health Technology Assessment dossiers. The app provides a collaborative space for multiple stakeholders over the course of request collection, dataset and analysis specification, output generation (using both SAS® and R), result delivery and change request.

Necessary metadata are collected with version tracking to ensure reproducibility and can be visualized as a traceability association map. A user-friendly interface is designed to guide the users and therefore increase the quality of inputs. An important aspect of the design is that any information should only be entered once, which leads to increasing efficiency and consistency while reducing human error.

INTRODUCTION

Health Technology Assessment (HTA) is a systematic evaluation of the clinical, economic, and social impacts of health technologies. It is often used by decision-makers to inform policy and reimbursement decisions for new drugs or medical devices. One of the key components of HTA is economic evaluation, which compares the costs and outcomes of different health technologies, and often requires additional clinical statistics beyond those reported in the clinical study report (CSR).

Creation of these outputs is a complex task. It involves multiple steps and cross-functional stakeholders supporting request generation, dataset preparation, specification documentation and output creation. Various types of programming software, such as SAS® and R are used to perform various statistical methods. Reproducibility and traceability of the outputs is an essential requirement in the process, which means that the same results can be obtained by executing the same code with the same data and parameters, and that the provenance of the outputs can always be traced back to the original sources.

PROCESS BEFORE H2H PORTAL

Before the H2H portal app was developed, the following process was in place:

1. Customers specified in an Excel document which outputs are required (Request).
2. The requests were copied into multiple Excel-based process documents (project plan and the list of Table, Listing & Figure) in slightly different formats.
3. The information was also translated into various technical documents and programs:
    a. Program specifications
    b. Table Mockup
    c. SAS program code for datasets
    d. SAS program code that generates outputs
4. The SAS programs were used to create the outputs.

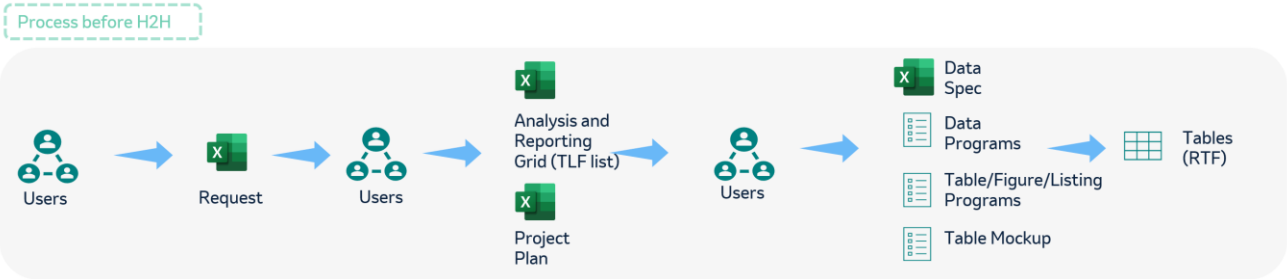

Figure 1: Process before H2H Portal

This complex process had several limitations. It relied heavily on repeated manual entry to document the same information across documents and platforms. This is suboptimal and may introduce human error and inconsistency.

As the steps are independent from each other, there was no systematic way to track and document the changes and decisions throughout the process. This reduces the transparency, reproducibility, and traceability of the analysis results.

The communication among stakeholders through these documents was facilitated primarily via email. This may lead to delays, misunderstandings, or conflicts.

USE CASES

H2H portal is an R Shiny app for end-to-end automation of outputs used in economic evaluation for HTA dossiers. It aims to achieve the following objectives:

- Provide a collaborative space for multiple stakeholders to communicate and coordinate over the course of request generation, dataset and analysis specification, output generation, result delivery, and change request.
- Support reproducible and traceable analyses and reporting from user requests to result generation. To achieve this goal, metadata is collected with version tracking.
- Integrate SAS® and R for output generation and delivery, using standardized templates and formats.

The end-to-end automation enables:

- Easy completion of project requests with version tracking
- Automated creation of key Analysis and Reporting (A&R) process documents and Mockup table package
- Tracking of relevant project metadata to enable automation
- Support of standardized as well as project specific Table, Listing & Figure (TLF)
- Target Users: Economic Modeler, HTA Statisticians and Statistical Programmers

The business drivers for this application are:

- to move away from Excel based request process
- to leverage ongoing standardization effort for majority of A&R deliverables
- to streamline project processes and documentation
- to efficiently produce large volume of outputs with high quality in a limited amount of time
- to provide an example of R based analysis & reporting (A&R) and
- to provide an example of end-to-end A&R automation using the Shiny framework.

In this paper, we will describe the high-level design, features and functions of the application, and its benefits and limitations.

MAIN DESIGN

The design of H2H Portal follows 3 key principles:

PRINCIPLE 1: END-TO-END

The app aims to cover from end-to-end all steps required in the analysis and reporting process. The following components have been successfully integrated into the app:

- Project Request
- Project Plan
- Mock table package
- Analysis Dataset specification
- A&R specification
- Table/Listing/Figure generation
- A&R process documentation

PRINCIPLE 2: AUTOMATION

- **Single-Entry**: The system enables the acquisition of necessary information at the right time and automatically propagates it across the lifecycle of the project analysis and reporting. This eliminates the need for redundant data entry and ensures that information is entered only once and consistently maintained throughout the process.

- **Guided Entry**: The friendly user interface (UI) of the portal is designed to guide the information entry with checks for quality inputs. The system also ensures consistency across projects by providing standard templates.

- **Auto-generation**: The system facilitates automation by enabling auto-generation of various key analysis and reporting deliverables including the calling programs, the analysis TLFs as well as necessary process documents, based on users' inputs and verification.

PRINCIPLE 3: TRACEABILITY AND REPRODUCIBILITY

The H2H Portal focuses on two critical aspects of data analysis: traceability and reproducibility, which enhance the reliability and transparency of the analytical process.

Traceability:
The app tracks and records all necessary metadata, including links between different components such as analysis datasets, study details, and function calls. This allows users to understand the provenance of the outputs in a comprehensive and reliable manner.

Reproducibility:
The H2H Portal ensures that the same results can be obtained consistently by running the same code with the same parameters. This is essential for validating and verifying analyses. Through versioning and tracking of relevant project metadata in the application, the analyses can be replicated precisely at any given time.

By incorporating traceability and reproducibility features, the H2H Portal ensures that the analytical process is robust, transparent, and consistent. This is particularly important in regulated industries where adherence to data integrity and auditability is crucial. Users can confidently reproduce results and trace back the origins of analyses, providing greater confidence in the accuracy and reliability of the outputs.

WORKFLOWS

The app consists of two main workflows (figure 2). The process starts with the request workflow, where the requestors specify all the analysis needs for a study. The information then flows to the Analysis & Reporting (A&R) workflow where the project Statisticians and programmers enter additional inputs, verify the outputs to complete the preparation of TLFs and process documents.

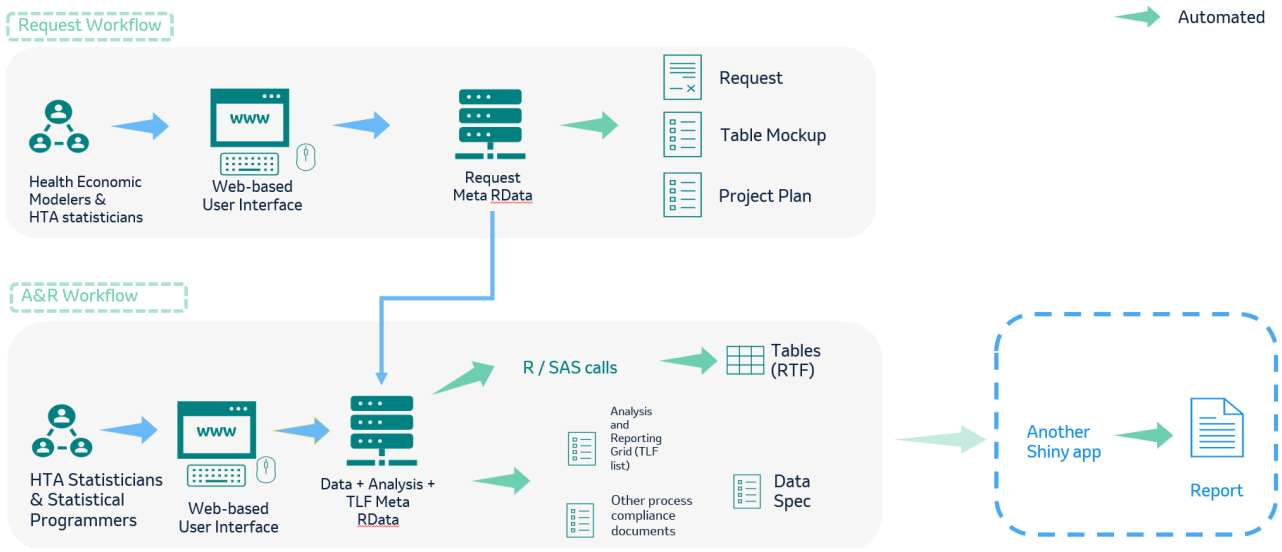

Figure 2: The full process covered by H2H Portal

H2H portal captures and preserves metadata at every stage of the process.

In the request workflow, the project Health Economic Modeler and/or HTA Statistician provides all the information about the analysis requests. This contains for example:

- General Study Info (Protocol, Data Cutoff Date, …)
- PICO (Population, Intervention, Comparator, Outcome) definition
    - Subpopulations
    - Treatment and Comparators
    - Requested analysis TLFs and datasets
- Timelines

This information is saved in the request metadata as an R object. During the A&R workflow initiation by the project statistician and programmer, the request metadata is uploaded to enable the request information to flow into the A&R metadata.

In the A&R workflow, users enter additional information that was not collected during the request stage:

- Data Meta - Analysis Datasets and specification
- Analysis Meta
    - PICOs derivation
    - Analysis methods
    - Programming assets
- TLF Meta - TLF layout, title, footers, filename…

The portal also provides some features to help the project team to review the outputs, and to support the promotion and generation of outputs from the test environment to the production environment.

STANDARD TEMPLATES

Multiple standard templates are integrated to improve deliverable consistency across studies (figure 3).

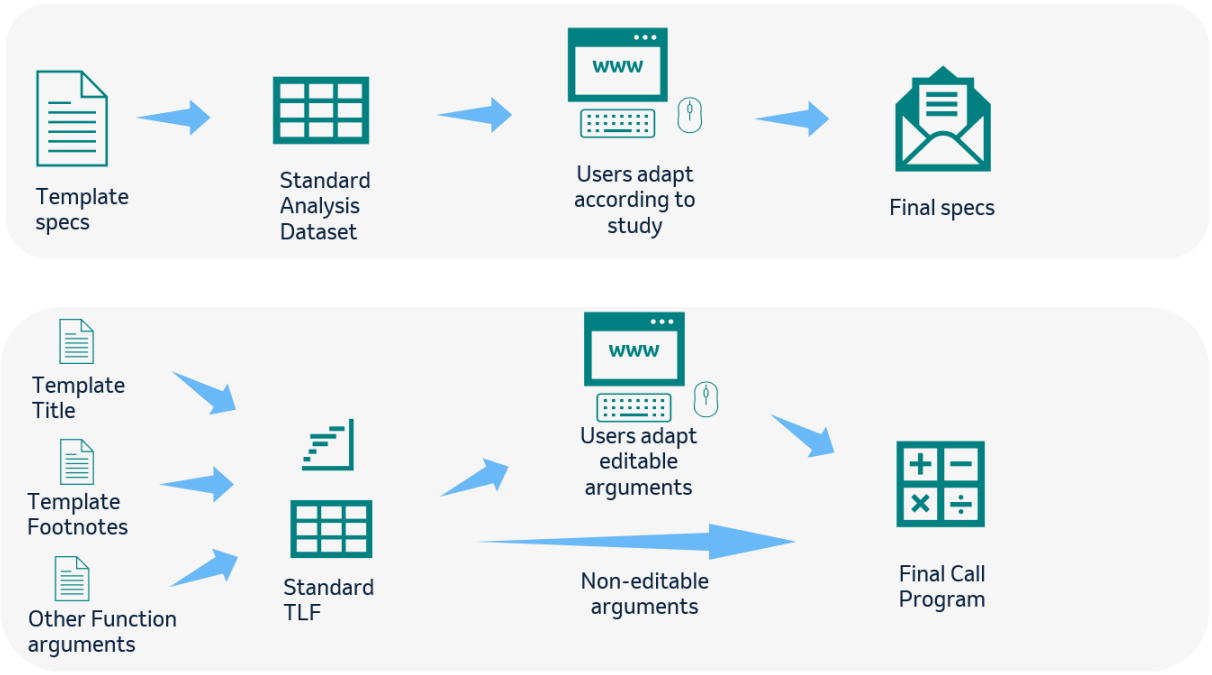

Figure 3: Use standard templates to facilitate automation

In the H2H Portal app, standard templates are provided to users for analysis and reporting deliverables. These templates serve as a starting point that can be modified and customized according to the study specific needs.

The use of standard templates in the H2H Portal app offers several advantages:

- Consistency: Standard templates ensure consistency across projects by providing pre-defined structures, formats, and elements. This helps to maintain a uniform approach to analysis and reporting, reducing the risk of errors and inconsistencies.
- Quality assurance: Standard templates are designed based on best practices and regulatory guidelines, ensuring that the outputs meet the standards and requirements of HTA agencies. This helps to improve the quality of analysis and reporting.
- Efficiency: By providing ready-to-use templates, the app saves time and effort for users. They do not need to start from scratch every time but can build upon the existing templates, thus accelerating the analysis and reporting process. Users can focus on the specific modifications and customization required for a particular study, rather than starting from a blank slate.

The templates provided in the H2H Portal app cover various aspects of analysis and reporting, including analysis datasets, TLFs. The templates include pre-defined titles, footnotes, and function arguments, which can be edited or modified as necessary. This allows users to tailor the templates to their specific study requirements while ensuring a consistent structure and format.

TRACEABILITY MAP

All the connections between different analysis and reporting components are captured and stored in the H2H metadata. The app provides a comprehensive overview of connections between datasets, specifications, programs, and other assets. A traceability map (figure 4) allows the users to visualize the links between various items and facilitates the identification of deliverables that haven't been refreshed despite changes in inputs. This not only aids the users to improve output accuracy and consistency, but also enhances reproducibility, traceability, and efficiency overall.

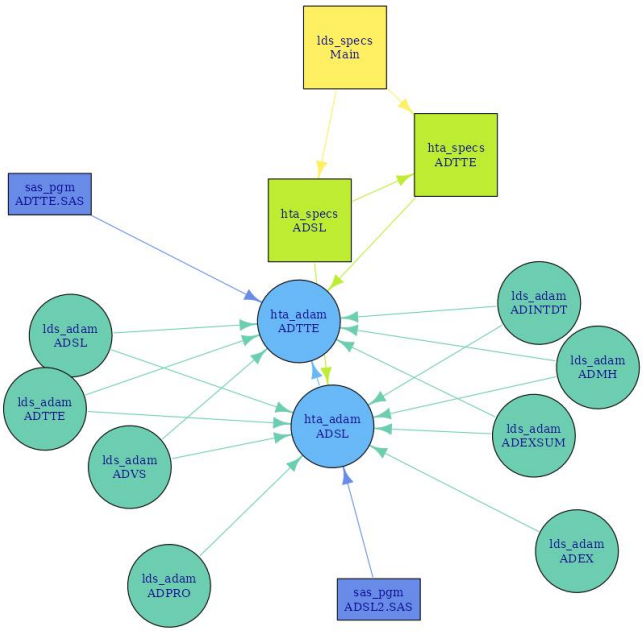

Figure 4: Traceability Map Visualization

CONCLUSION

The H2H portal app is an innovative and useful tool that automates many steps in the generation of clinical statistics deliverables for HTA dossiers. The app provides a collaborative space for multiple stakeholders to communicate and coordinate over the course of request workflow.

Several challenges have been observed during the development of the H2H Portal.

The app required integration of multiple existing sub-processes, e.g. project request, dataset specification, analysis specification, and output generation, each developed and owned by different team. Pro-active alignment and smart design were crucial to ensure end-to-end automation with a seamless and streamlined workflow.

Finding the right balance between simplicity and functionality was another challenge. The app must be user-friendly and intuitive, guiding users through the process with improved accuracy. At the same time, it should contain the necessary flexibility and capabilities to cover the full A&R process required across all projects.

It was also a challenge to come up with the most effective way to modularize the design, and to distribute the work among multiple individuals involved in the development process. This required coordination and collaboration to ensure that each component of the app was developed and integrated seamlessly.

EXPECTED RESULTS

The H2H Portal app aims to enhance the A&R process to generate additional clinical statistics beyond the CSR in support of economic evaluation in Health Technology Assessment (HTA) dossiers. Once released, the following outcomes are expected:

- Enhanced Quality: The app ensures the completeness and correctness of inputs by guiding users through the process with quality checks, which leads to improved output consistency and enhanced deliverable quality.
- Improved Efficiency: The H2H Portal app streamlines the workflow by automating various steps and eliminating manual entry and transfer of information. As a result, it reduces the time and effort for the project team and improves the overall efficiency of the process.

- Increased Reproducibility: The app collects metadata and maintains version tracking, ensuring the reproducibility of results. This allows the same results to be obtained consistently by running the same code with the same parameters, providing confidence in the validity of the analysis.

- Improved Collaboration: The H2H Portal app provides a collaborative space for multiple stakeholders involved in the process. It facilitates communication and coordination, leading to better collaboration with less delays, misunderstandings, and conflicts.

- Comprehensive Traceability: The traceability association map provided by the app offers a clear overview of the information flow and relationships between different components within the A&R workflow. It assists the users to visualize the provenance and lineage of the outputs and enhances the process transparency.

- Scalability and Adaptability: The H2H Portal app provides capability to handle various projects and adapt to evolving needs. It showcases its potential for future enhancements and scalability to accommodate different analyses and reporting requirements.

FUTURE DEVELOPMENT

The development of the H2H Portal app does not stop at its initial release. The following areas of future development have been identified:

1. Integrate User Feedback: Gathering user feedback is crucial for continuous improvement. By soliciting input and insights from users, the H2H Portal app can identify areas for enhancement and address any pain points or usability issues. Incorporating user feedback will result in a more refined and user-friendly application.

2. Comparison Module: Implementing a module that allows for the comparison between different versions of items, such as dataset specifications and datasets, will be a valuable addition. This module would enable users to visualize and assess the impact of changes between versions, enhancing decision-making and improving consistency and accuracy in the analysis process.

3. Cross-study metadata reporting: The metadata collected via H2H provides valuable data for further process improvements across projects. A possible use case will be to identify analyses requested by multiple projects as potential candidates for further standardization. Another example is to produce metrics dashboard within/across projects on efficiency/workload.

By focusing on these future development areas, the H2H Portal app will continue to evolve and adapt to the changing needs and requirements of the economic evaluation for HTA dossiers process. The integration of user feedback, implementation of a comparison module, and the addition of more automated features will further enhance the efficiency, accuracy, and user experience of the app.

We believe that the H2H portal app is a valuable addition to our statistical and statistical programming toolbox that can help us deliver high-quality statistical outputs for HTA dossiers in a timely manner. We hope that our experience can inspire other R Shiny developers to continue to implement innovative solution to address business need.

GLOSSARY

A&R	Analysis and Reporting
CSR	Clinical Study Report
H2H	Health Economic modelers <-> HTA statisticians/programmers
HTA	Health Technology Assessment
LDS	Late Development Statistics
PICO	Population, Intervention, Comparator, Outcome
RTF	Rich Text Format
TLF	Table, Listing, Figure
UI	User Interface

ACKNOWLEDGMENTS

We would like to thank the development team for their outstanding dedication and effort in creating the H2H Portal app. The team's hard work and innovative thinking have been instrumental in the success of this project. We are grateful for their expertise and commitment to delivering a high-quality automation tool.

Additionally, we would like to acknowledge the validation team for their meticulous testing and validation of the H2H Portal app. Their attention to detail has ensured the reliability and accuracy of the app.

Moreover, we would also like to thank our sponsors, business stakeholders, and everyone who has supported the development of this application.

CONTACT INFORMATION

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

Julian Schmocker
julian.benjamin.schmocker@msd.com

Qian Wang
qian.wang1@msd.com

X. Gregory Chen
xiangyi.gregory.chen@msd.com