PharmaSUG China 2025 - Paper AA-115

Automating IRC Data Validation: From Scripted Code to Configurable Rules via an Extensible Platform

Wenjun Zhu, Jiangsu Hengrui Pharmaceuticals Co., Ltd.

ABSTRACT

The evolution of biometrics in clinical trials is driven by synergistic advances in regulatory science and computational technologies. Previously, the adoption of electronic Case Report Forms (eCRFs) rendered a kind of workforce obsolete, and now, automation is being deployed across various work modules in clinical trials, assisting in problem-solving and reducing manual work.

Blinded Independent Review Committee (BIRC) data governance serves as an illustrative case. As a key endpoint in new drug applications, BIRC data requires systematic quality control. Despite vendors' established imaging review workflows, sponsors face challenges in verifying quality compliance of data mapping results due to different CDISC (Clinical Data Interchange Standards Consortium) implementation approaches across vendors. This context makes the automated validation system developed by CDISC-specialized programmers particularly valuable, delivering practical solutions that enhance endpoint reliability while facilitating the integration of data management and statistical analysis.

Our proposed system transforms the traditional three-party collaboration process into a unified workflow. We've developed a systematic framework that leverages configurable rule sets and modular logic components to automate validation processes. While our paper primarily focuses on the automation framework's architecture and rule execution—serving as the critical bridge between CDISC standards and validation practices—the system demonstrably supports extensibility. Leveraging our automation-ready rule setting framework, we can seamlessly integrate AI assistants to enable natural language data validation. Critically, this empowers vertical clinical analysis tools to resonate within the AI ecosystem—amplifying their utility through orchestrated interoperability. By bridging CDISC standards with automated applications, our framework offers a scalable solution for endpoint reliability and streamlines collaboration between data management and statistical analysis teams.

INTRODUCTION

To bridge the gap between data management and statistical analysis standardization, biometrics professionals must navigate a complex environment requiring simultaneous expertise in radiological assessment principles and CDISC SDTM (Study Data Tabulation Model) specifications, while also coordinating with sponsor requirements. This multi-way coordination process is inherently challenging, but what if we could streamline this into a unified, efficient automated workflow through a systematic framework approach?

Our implementation approach is to structured in three sequential phases as below: Criteria interpretation, the codifying of the rule sets, and UI (User Interface) implementation.

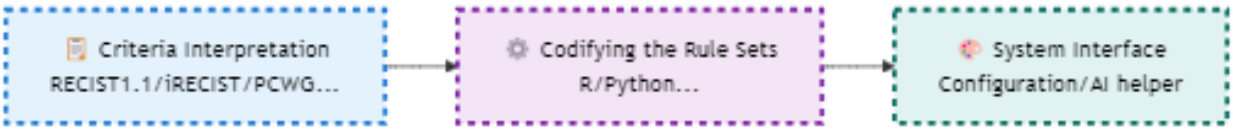

Figure 1. Workflow in three phases

The **first step** involves comprehensive interpretation of oncology assessment criteria (RECIST 1.1 (Response Evaluation Criteria in Solid Tumors), iRECIST (immune Response Evaluation Criteria in Solid Tumors), PCWG3 (Prostate Cancer Working Group), etc.) to extract validation rules. The **second step** focuses on program transformation, integrating CDISC SDTM standards with automated validation logic to create executable code modules. The **final step** encompasses system interface implementation.

Next, the implementation of the above plan will be elaborated.

STEP 1: CRITERIA INTERPRETATION

WHY RESPONSE CRITERIA BREAKDOWN?

While professional institutions and organizations provide systematic assessment frameworks, such as RECIST 1.1 and PCWG3, transitioning actual assessment into standardized, analysis-ready data remains challenging. To ensuring consistency and accuracy of the data, the validation demands extracting clear, logical, and structured rules from these guidelines.

Example of a Logic Conflict

Here is an example, according to RECIST 1.1, the Overall Response must be determined based on the measurement data of both target and non-target lesions. If the overall response is CR (Complete Response), it requires the dual conditions that all target lesions have disappeared (CR) and all non-target lesions show no disease, while also excluding the presence of new lesions. That means if a time point has PR in target lesions, the Overall Response could not be CR – if evidence is found, a potential logic error is detected.

Understanding these interdependent mechanisms among data points allows us to further develop rule sets based on guidelines and standards.

Other Logics Underlying

It is worth noting that rigorous BIRC tumor assessments often require evaluation by more than one reviewer, and sometimes arbitration, which further increases the complexity of data logic checking. This means that the automated system must not only verify whether the measurements at a single time point meet the standards, but also track dynamic consistency across time points (such as confirming whether new lesions persist in subsequent assessments) and evaluators. In some special cases that violate logical rules, the system should also check the reviewers' comments (for example, if a reviewer changes an assessment but does not retrospectively correct the previous “wrong” assessment, only leaving a comment as an explanation).

Analytical and Clinical Expertise Required

These rules, both simple and complex, require the involvement of statistical analysis experts to sort out and summarize rules that are both efficient to implement and precise and rigorous. At the same time, we can separate out some violations that cannot be structurally processed, leaving them for human experts to judge their reasonableness.

Based on the above understanding, we have extracted rule sets. The next step is to organize them into program-readable, structured rule set libraries.

Rule-sets Extracted

The JSON structured content below presents an example for solid tumor evaluations based on RECIST 1.1, the abbreviations are from the controlled terminology in the SDTM IG (Implementation Guidance) or commonly used internal abbreviations, for example, TLs stands for target lesions; NTLs stands for non-target lesions:


```
    {
      "category": "RECIST1.1",
      "subCategory": "Responses",
      "ruleId": "R001",
      "sources": ["rs","tr"],
      "MSG_CN": "当总评结果为 PD 时，该访视靶病灶或/和非靶病灶结果不为 PD，或新病灶不为
UNEQUIVOCAL/Y。首次未明确的新病灶(Equivocal)可由下次明确(Unequivocal)的新病灶确认，此时总评的 PD 会回
溯性更改为上次访视的日期",
      "MSG_EN": "The OVRLRESP is PD but the TL and/or NTL is not PD, and the new lesions is
not UNEQUIVOCAL/Y for this visit, the previous visit's equivocal new lesion can be confirmed
by a subsequent unequivocal new lesion. In this case, the date of PD will be changed to the
date of the previous visit"
```

```
}
```

Another JSON structured paragraph below contains an example for the PCWG3 criteria, the abbreviations are from the controlled terminology in the SDTM IG (Implementation Guidance) or commonly used internal abbreviations, for example, Pdu stands for PD unconfirmed:

```
{
  "category": "PCWG3",
  "subCategory": "Dates",
  "ruleId": "D003",
  "sources": ["rs"],
  "MSG_CN": "只有一次访视且为 PD 而不是 PDu ",
  "MSG_EN": " Only one visit and the result is PD, not PDu"
}
```

MACHINE-READABLE RULE SPECIFICATION

The following is an example of a machine-parseable specification written in JSON format. This rule is a complex response validation check from RECIST 1.1 (Rule-sets Extracted), described as: "The OVRLRESP is PD but the TL and/or NTL is not PD, and the new lesions is not UNEQUIVOCAL/Y for this visit, the previous visit's equivocal new lesion can be confirmed by a subsequent unequivocal new lesion." Here, "sources" indicates that the data comes from both RS and TR domains, these source data will be loaded and cross-dataset validation will be performed. The rule involves complex logic including multiple condition checks, cross-visit validation, and equivocal lesion confirming mechanisms.

Below is the example specification in JSON format:

```
{
  "ruleId": "R001",
  "description": {
    "cn": "当总评结果为 PD 时，该访视靶病灶或/和非靶病灶结果不为 PD，或新病灶不为 UNEQUIVOCAL。首次未明确的新病灶……",
    "en": "The OVRLRESP is PD but the TL and/or NTL is not PD, and the new lesions is not UNEQUIVOCAL……"
  },
  "sources": ["rs", "tr"],
  "conditions": {
    "filter": {
      "RSCAT": "RECIST1.1",
      "RSTESTCD": "OVRLRESP",
      "RSSTRESC": "PD"
    },
    "check": {
      "type": "complex_response_validation",
      "logic": {
        "overall_pd": true,
        "target_not_pd": "!has_pd_target",
        "nontarget_not_pd": "!has_pd_nontarget",
        "new_lesion_not_unequivocal": "!has_newlesion_unequivocal",
        "equivocal_confirmation": "check_equivocal_confirmation"
      },
      "groupBy": ["SUBJECTID", "VISIT"],
      "crossVisitCheck": {
        "enabled": true,
        "type": "equivocal_confirmation",
        "previousVisit": "check_equivocal_new_lesion",
        "currentVisit": "confirm_unequivocal_new_lesion"
      }
    }
  }
}
```

STEP 2: R SHINY – RULE ENGINE SELECTION

While SAS remains the dominant language for clinical trial programmers, an increasing number of professionals are adopting R due to its powerful data manipulation capabilities, rich statistical ecosystem, and open-source nature. In this project, R Shiny emerged as the optimal technology stack for building interactive validation systems.

R SHINY'S WORKING MECHANISM

R Shiny's reactive programming model automatically handles dependencies between user inputs and outputs, making it ideal for dynamic rule execution based on user configurations and data inputs.

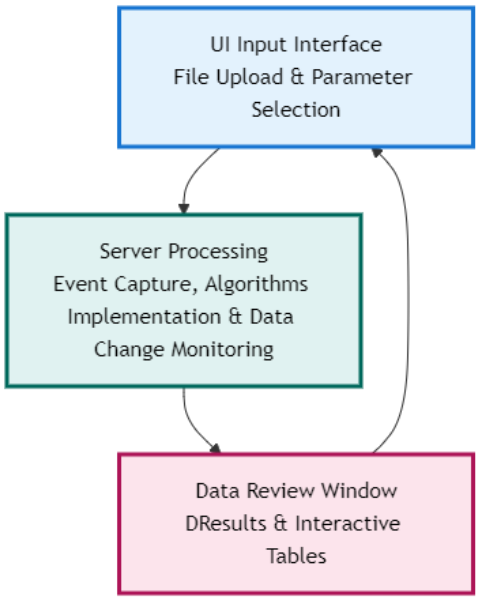

Figure 2. R Shiny's working mechanism

R Shiny's interaction mechanism can be organized into three core functional blocks:

1. **UI Input Interface**: Implements HTML form elements and interactive widgets for file uploads, dropdown selections, and button controls - the client-side input layer that captures user interactions through web browser events and transmits them to the server.
2. **Server Processing**: Handles server-side event listening, processes incoming HTTP requests, executes computational algorithms, and maintains reactive state management through a web socket connection that enables real-time bidirectional communication between client and server.
3. **Data Review Window**: Also a part of UI, renders dynamic HTML content including interactive data tables with sorting and filtering capabilities, statistical charts and plots.

PLATFORM'S ARCHITECTURE

This simplified circular mechanism ensures immediate response to user operations and real-time data updates, making complex validation logic accessible through an intuitive interface. Hence, the developed R Shiny application adopts a modular architecture with clear separation of concerns:

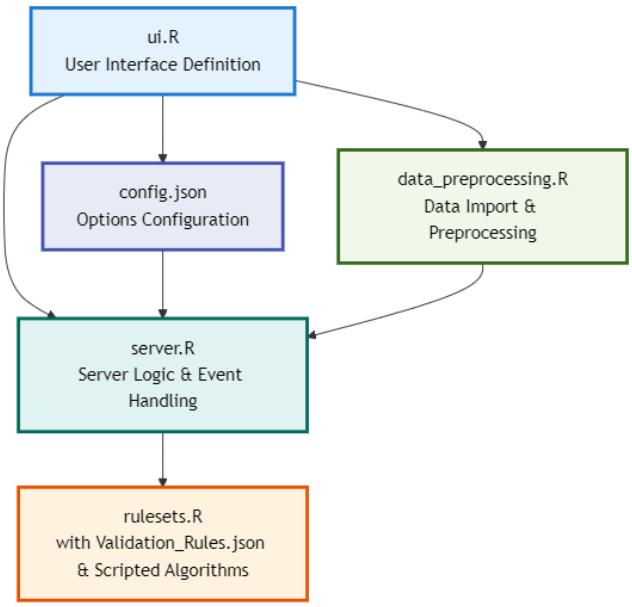

Figure 3. Modular architecture

This modular structure enables:

- **Separation of concern**: Each module handles specific functionality

- **Maintainability**: Easy to update individual components without affecting others

- **Reusability**: Rules and configurations can be shared across different applications

- **Scalability**: New validation rules can be added by updating the library of rule sets(validation_rules.json) and the algorithms.

CODIFIED EXAMPLE

To illustrate the rule implementation approach, here is a very simple example to demonstrate how rules can be efficiently extended and maintained within the platform:

Rule D002: "Has evaluation result but no evaluation date"

This rule checks for records where there is an evaluation result (RSORRES) but the evaluation date (RSDTC) is missing. The algorithm is straightforward but demonstrates the core validation logic:

```
check_D002 <- function(rs_data) {
  results <- rs_data %>%
    mutate(
      RSCAT = gsub(" ", "", toupper(trimws(RSCAT)))  # Standardize category
labels from heterogeneous input data
    ) %>%
    filter(
      RSCAT == 'RECIST1.1',  # Filter for RECIST1.1 records
      !is.na(RSORRES) & is.na(RSDTC)  # Has result but no date
    ) %>%
    select(SUBJECTID, VISIT, RSDTC, RSORRES) %>%
    rename(DATEF = RSDTC) %>%
    mutate(Domain = "RS")
  return(results)
}
```

By developing such rules, you can progressively enhance the rule set library to accommodate evolving validation requirements while preserving an executable and intuitive structure. Operating within R Shiny's reactive framework, these rules dynamically execute validation processes triggered by user interactions, with outcomes determined by configured parameters and data inputs.

STEP 3: FUNCTIONAL DESIGN AND THE WORKFLOW

RATIONALE

The platform's user interface features a two-panel structure: a left sidebar for data input/configuration, a tabbed main content area for results display with a responsive layout, which is a very intuitive and developer-friendly R Shiny application design.

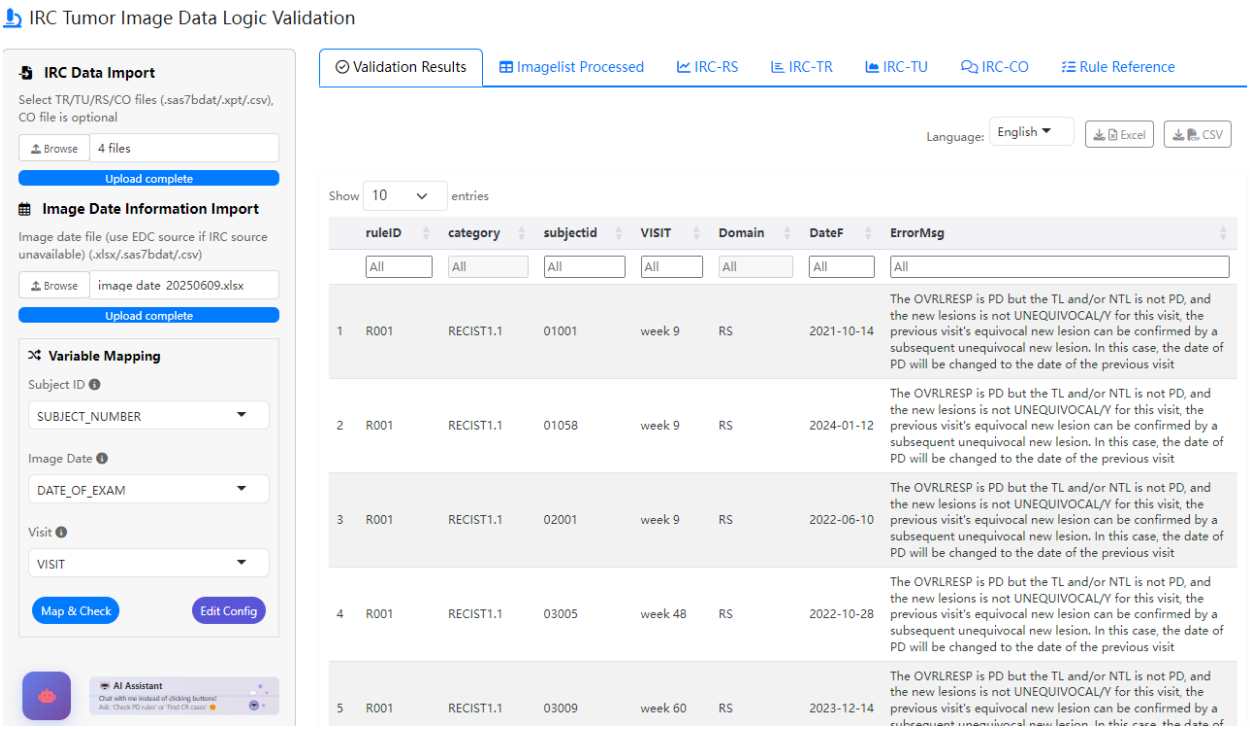

Figure 4. The UI of the application

The selection of the technical stack – specifically, why R Shiny was chosen for platform development – has already been explained in the previous section **(Step 2: R Shiny – Rule Engine Selection).** However, ensuring the effective execution of these rules requires that the quality and structure of input data meet expectations. Therefore, integrating domain knowledge of CDISC-standard and establishing robust data preprocessing workflows are essential. Different vendors may vary slightly in their implementation of CDISC mapping logic, and may even introduce technical problems. The system must address these challenges and provide a reliable alternative solution for subsequent logic validation. These considerations and design philosophies led us to modularize the functionality into the following components:

- Data Upload Module
- Warning and Preprocessing Function
- Configuration Interface
- Results Display Window

The next section will detail the design princple and functions of these interfaces.

DATA UPLOAD AND PROCESSING

This data module handles two critical data input streams:

IRC Data Import: This component accepts multiple file formats (.sas7bdat, .xpt, .csv) for the essential SDTM domains: TR (Tumor Results), TU (Tumor Identification), RS (Disease Response and Clinical Classification) with optional CO (Comments) domain support. The interface automatically detects file prefixes and categorizes them accordingly, eliminating manual file sorting requirements.

Image Scan Date Import/Mapping: This specialized upload section handles imaging acquisition date data, which is crucial for validation logic. The interface supports multiple source formats

(.xlsx, .sas7bdat, .csv) and provides fallback options when IRC source data is unavailable, allowing users to utilize EDC-derived imaging dates as alternatives. Users could select different variable from drop-list if the variable names from vendors are atypical.

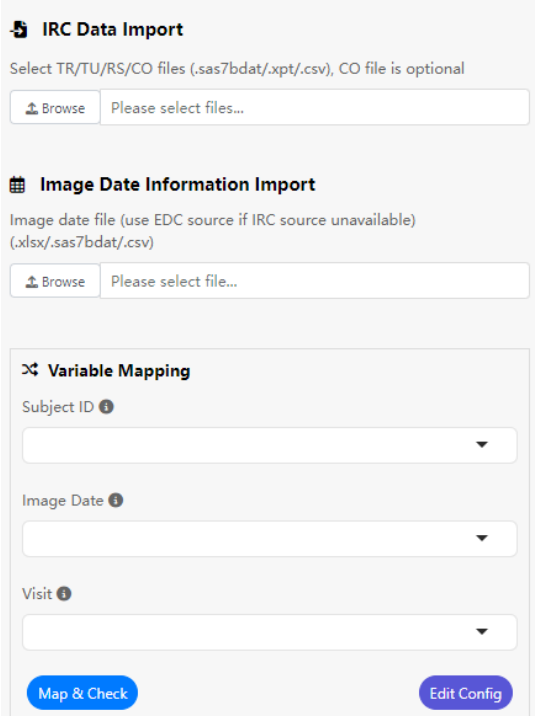

Figure 5. Data uploading and processing panel

The functional part of processing the **Image Scan Date** may appear unconventional to many, if not all, clinical trial analysts not familiar with oncology efficacy response analysis, but it is essential for response dates' cross-validation. In response analysis related to solid tumors, the date decision of PD or Non-PD responses rely not only on lesion measurement dates recorded in the TR domain, but also requires the inclusion of acquisition dates of all scanned images for all visits (including images determined as "no lesions"), which are usually provided by the vendors as a seperated image list, because TR domain usually collects "measurable lesions" only. Especially in certain tumor types such as prostate cancer, missing such date information often leads to inconsistencies in efficacy date logic. The data needed for cross-validation is simple in structure (containing SUBJECTID, IMAGEDATE, VISIT), but the raw data provided by vendors often differ significantly in date formats and visit naming. The system addresses these heterogeneities through preset, configurable data processing logic (such as complex date parsing rules and visit mapping mechanisms).

WARNING & PREPROCESSING MODULE

The platform also implements a warning mechanism that mainly operates in the data preprocessing part. This part is separated from the core validation rules part, based on the principle that necessary error prevention in preprocessing ensures criteria-based validation rules can be effectively applied. These two layers address distinct aspects, core validation rules primarily impact critical study outcomes like primary endpoints; while preprocessing warnings focus on data quality and formatting.

Below is an example of a warning log.

CheckDate	Domain	CheckType	Message
2025-07-05	CO	id_var	No USUBJID variable found, using SUBJID instead
2025-07-05	RS	id_var	No USUBJID variable found, using SUBJID instead
2025-07-05	RS	empty_acptfl	Subjects with all empty RSACPTFL: CN001001, CN001005, CN001007 and more
2025-07-05	RS	adjundication	CN001008 at Week 12 has: R1 : SD; R2 : PR but no adjundication

Figure 6 Warning log screenshot

When users upload the data, the application performs extensive data standardization operations, generating detailed warning logs for any detected issues.

The warning log captures four key information categories:

Timestamp for chronological tracking

Dataset identification for source attribution

Check type classification for issue categorization

Detailed message **content** for user guidance

CONFIGURATION SETUP

The configuration setup window represents the application's most sophisticated user interaction component, providing granular control over validation parameters and processing options. Here is how it looks like:

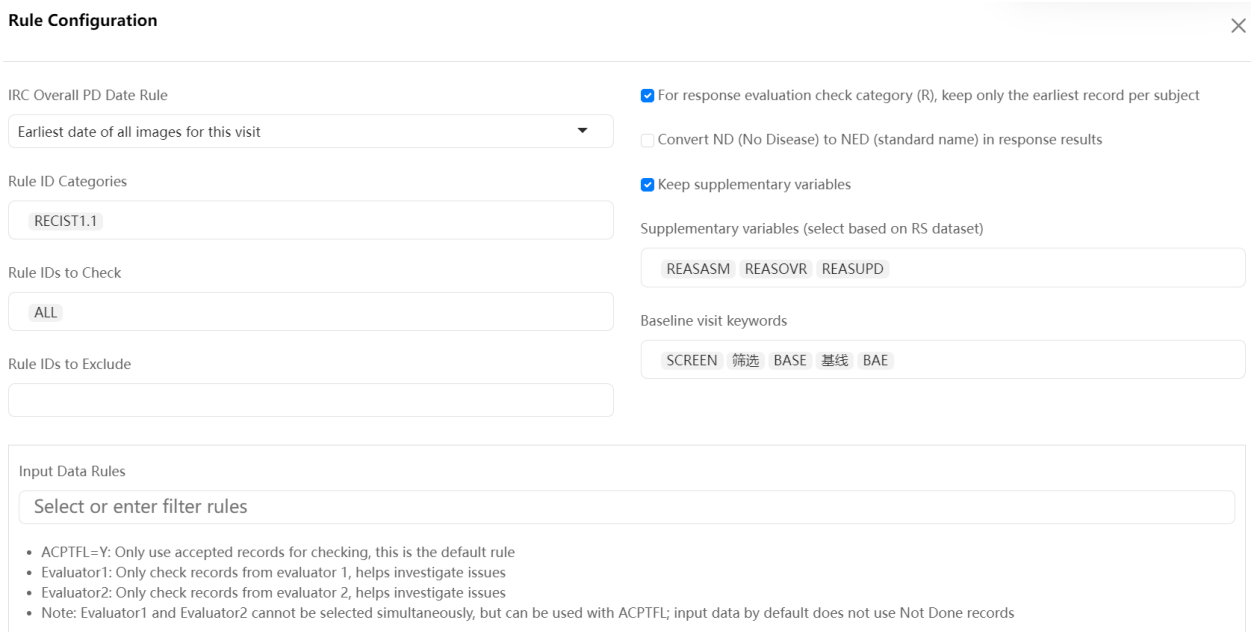

Figure 7. Configuration setup window

The modal-based interface is organized into logical sections, it allows users to select IRC Overall PD Date Rules from two methodologies ("Earliest date of all images for this visit" or "Earliest date of PD assessment for this visit") and customize baseline visit keywords to align with study-specific terminology. Additionally, it supports subset filtering (e.g., Evaluator 1 only, Evaluator 2 only, or accepted records only) along with options for non-duplicated results and comments viewing. The Rule Set Management section enables category-based filtering (RECIST1.1, PCWG3, or comprehensive ALL selection), individual rule inclusion/exclusion for targeted validation, and configuration of rule priority and execution order.

The configuration interface incorporates intelligent defaults based on industry best practices while maintaining flexibility for study-specific requirements. The modal design ensures focused attention on configuration tasks while preserving access to the main interface.

RESULTS DISPLAY WINDOW

As a platform primarily designed for data analysis professionals, the Result Display interface is actually the most intuitive, with relatively simple features to facilitate users' browsing and inspection of data frames. The main content area employs a tabbed navigation system to organize diverse output categories efficiently. The display looks like this:

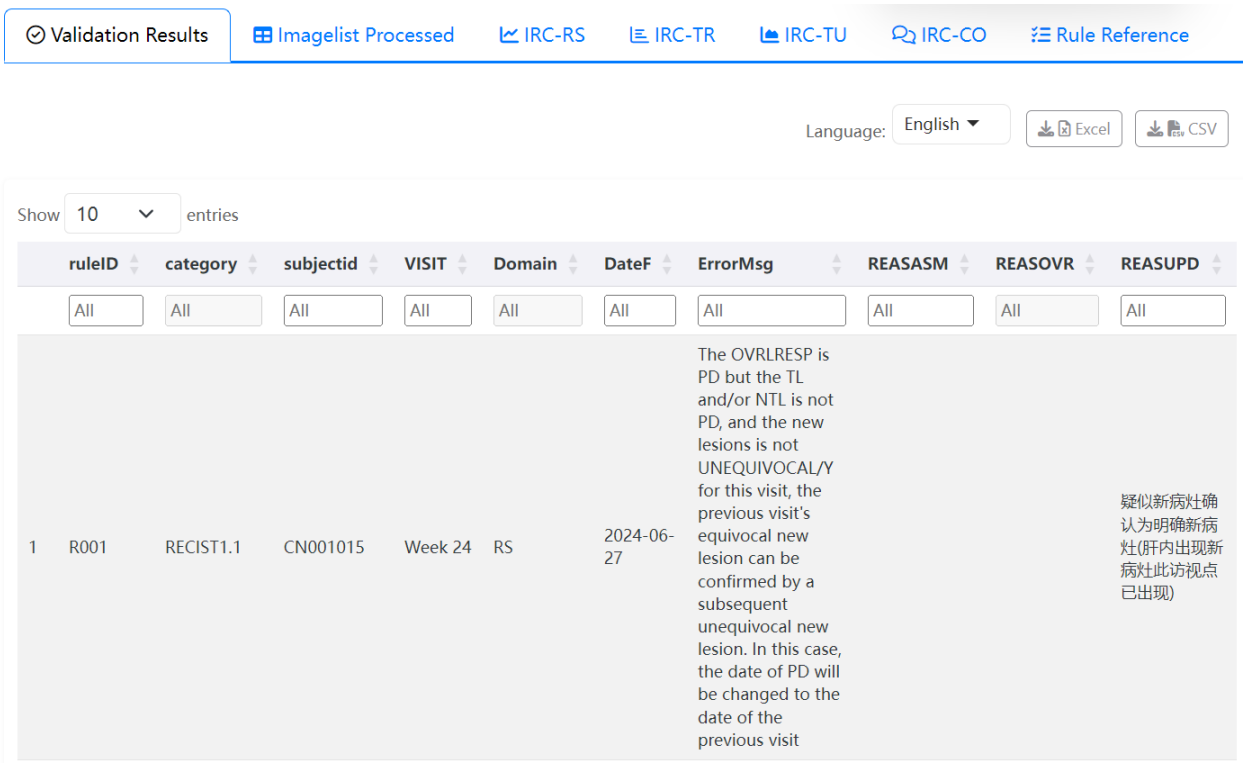

Figure 8. Validation results window

In addition to viewing and inspecting all uploaded raw data based on issues, this "Validation Results" window can present records violates the rules. In this example, variables of REASASM (Reason for Assessment), REASOVR (Reason for Override), REASUPD (Reason for Update) are also selectively incorporated, enabling users **–** beyond structural checks **–** to draw further conclusions based on the contextual insights provided in the comments.

ADD-ON: AI INTEGRATION

While the core validation framework operates on pre-configured rules (Step 1: Criteria Interpretation), an interesting interaction layer – a natural language interface powered by Large Language Model (LLM) is introduced to this automation platform. This optional component serves as an accessibility enhancer, allowing users to do easy validation checks via conversational queries through a user-friendly interface.

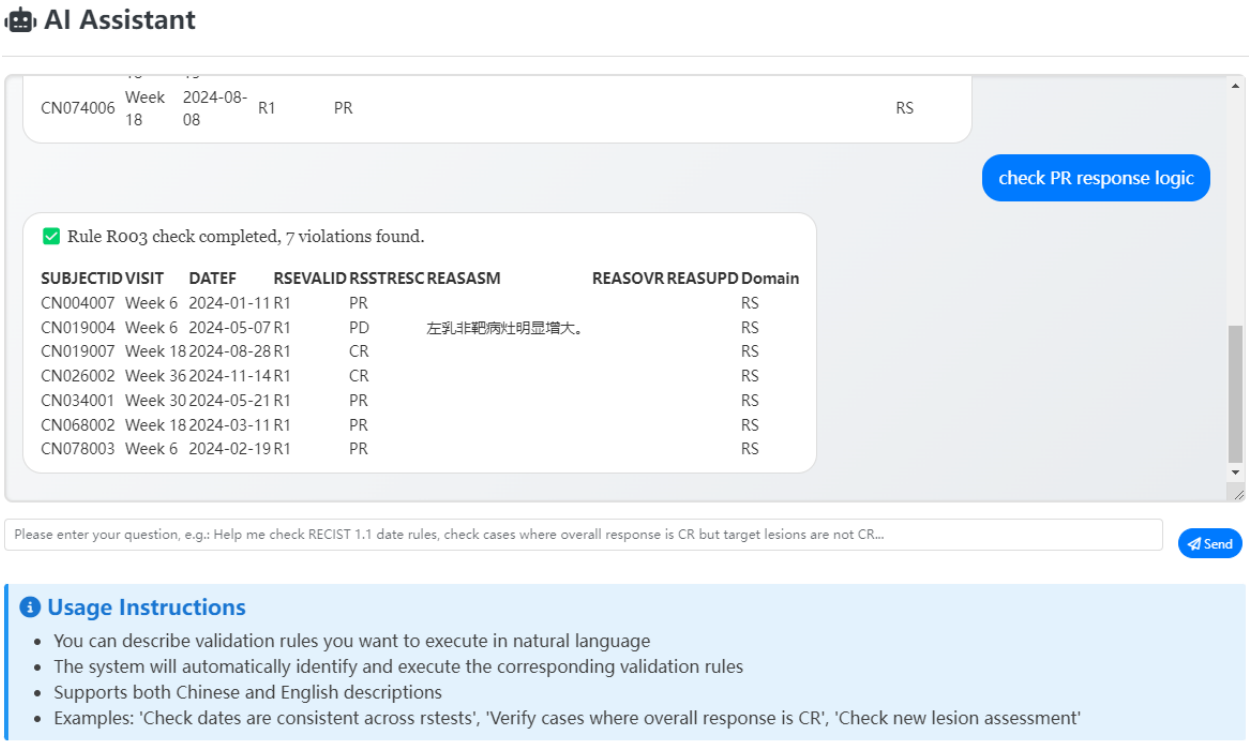

Figure 9. Natural language validation checks

The AI layer acts as a thin translation bridge, converting natural language requests (e.g., 'Check PD assessments without lesion progression') into predefined rule IDs. When users pose questions, the assistant performs the checks and returns tabular results, functionally similar to the platform's automated validation (since both are based on the same core engine). While this functionality itself is not a fundamental innovation, it extends our thinking toward broader applications:

This prototype demonstrates the feasibility of human-AI collaboration in clinical data validation. In many ways, it shares conceptual similarities with the widely-discussed Model Context Protocol (MCP): the AI chat interface essentially implements MCP Host functionality, receiving natural language instructions and converting them into calls to internal validation rules. However, it is not yet a true MCP tool, as the chat interface maintains tighter coupling with the underlying system compared to MCP's standardized decoupling philosophy. Moreover, this solution is designed for secure enterprise deployment within clinical research organizations, adhering to stringent data confidentiality standards, which conflicts with the MCP protocol’s documented limitations.

Nevertheless, this application provides a compact and inspiring example for enterprises to build internal standards enabling further AI-coordinated tool deployment. Subsequent explorations for broader applications for AI integration within enterprise environments could leverage JSON-based communication protocols and standardized instruction interfaces to invoke various internal tools and to empower enterprise personnel to solve problems with enhanced precision.

Such an approach could transform how organizations utilize their existing tool ecosystems through natural language interfaces, accelerating the creation of more accessible and efficient workflows for both domain experts and non-technical users alike. Applying this vision to clinical trial companies could optimize the alignment between current data governance practices and analytical operations.

CONCLUSION

The development of this automated IRC data validation platform demonstrates how systematic framework approaches can transform traditionally manual, error-prone processes into efficient, configurable, and intelligent workflows. Our work addresses a critical gap in the clinical trial ecosystem by bridging CDISC

standards-based results with complex evaluation logics with practical validation needs through breaking down the complex rules and transforming them into automated validation workflows.

While our automated IRC data validation system demonstrates significant potential, several limitations should be acknowledged. The current implementation is primarily focused on oncology assessment criteria (RECIST 1.1, iRECIST, PCWG3), and its applicability to other therapeutic areas or assessment frameworks remains to be validated.

Future enterprise implementations may explore deeper AI integration, demonstrating expanded human-AI collaboration frameworks – such as adopting AI to help human programmers codify and expand the rule libraries for specialized domains - while extending this validation mechanism to broader clinical trial workflows and adjacent applications.

REFERENCES

Eisenhauer et al. 2009. *"New Response Evaluation Criteria in Solid Tumours: Revised RECIST 1.1." Eur J Cancer* 45:228-247. DOI: 10.1016/j.ejca.2008.10.026.

Scher, H.I., M.J. Morris, W.M. Stadler, et al. 2015**.** *"The Prostate Cancer Working Group 3 (PCWG3) Consensus for Trials in Castration-Resistant Prostate Cancer (CRPC)*." *Journal of Clinical Oncology*, 33(15_suppl):5000.DOI: 10.1200/jco.2015.33.15_suppl.5000

U.S. FDA. 2015. "*Clinical Trial Endpoints for NSCLC Drug Approval: Guidance for Industry.*" Silver Spring, MD: FDA. Available at https://www.fda.gov/media/116860/download

ACKNOWLEDGMENTS

The authors wish to express profound gratitude to **Dr. Xianfeng Zhou** for her seminal contribution to her proposal of the IRC data validation projects, which led to originating this paper. Dr. Zhou's idea sparked this application and her insights continuously provided methodological guidance.

CONTACT INFORMATION

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

Wenjun Zhu
Jiangsu Hengrui Pharmaceuticals Co., Ltd.
86-17721285930
Wenjun.zhu@hengrui.com