

2026-04-30

R4SUB: Submission Readiness Scoring. A Standardized Evidence Schema and Scoring Framework for Multi-Indicator Readiness Assessment in R

Pawan Rama Mali 
R4SUB Project (independent open-source)

Aishwarya Amonkar
Techworkslab

Table of contents

Executive Summary	2
Signatories	3
Project team	3
Contributors	3
Consulted	3
The Problem	3
What the problem is	3
Who it affects	4
Why it is a problem	4
What solving it enables	4
Existing work and previous attempts	4
Why the R Consortium	5
The proposal	5
Overview	5
Detail	6
Project plan	8
Start-up phase	8
Technical delivery	8
Other aspects	9

Sustainability beyond the grant	10
Budget and funding plan	10
Success	11
Definition of done	11
Measuring success	11
Future work	11

Executive Summary

A regulatory affairs lead at a mid sized biotech is preparing a New Drug Application. She has Define-XML metadata, a Pinnacle 21 validation report, ADaM and SDTM datasets, and an Analysis Data Reviewer’s Guide. Internally, three programmers and one statistician have been working through a spreadsheet of “submission readiness” checks for two months. Each reviewer scores items differently. The team cannot tell whether the submission is ready, partially ready, or carries hidden gaps that will trigger an FDA Refuse to File. There is no shared way to compute, compare, or audit the answer.

The same problem appears in adjacent R communities. The R Validation Hub describes a structurally identical workflow when assessing whether an R package is qualified for use in a regulated environment. Reproducibility researchers face the same question when scoring whether a study artifact set meets open science criteria. Data quality teams ask it of every dataset that enters a pipeline. In each case the workflow is the same: define indicators, gather evidence, score each indicator, aggregate into dimension scores, and decide. The R ecosystem has no shared infrastructure for this work, so each community rebuilds it from scratch.

R4SUB is a standardized evidence schema and scoring framework that gives this class of problem a single, reusable foundation. Six packages already on CRAN cover the schema, scoring engine, regulatory profiles, and Shiny review dashboard. The framework is functional end to end on its first application (regulatory submissions for FDA, EMA, and PMDA), but the scoring weights are still based on expert opinion rather than empirical outcome data, and the schema is not yet formally specified for external adoption.

We request \$12,500 over six months (July to December 2026) for three milestones. Milestone 1 (\$3,500) formalizes the evidence schema as a versioned JSON Schema specification with a CRAN distributed validator. Milestone 2 (\$6,000) calibrates the scoring weights against twenty or more anonymized regulatory submissions with known outcomes, replacing expert estimates with evidence based weights. Milestone 3 (\$3,000) publishes integration vignettes with pharmaverse (admiral, xportr, riskmetric) and a worked non-pharma example showing how the same infrastructure applies to R package validation.

The work raises R4SUB from a functional prototype that one project uses to standardized infrastructure that any R community doing indicator based readiness assessment can adopt and extend.

Signatories

Project team

Pawan Rama Mali is the Project Lead and original author of the R4SUB ecosystem. He maintains all eight packages (six on CRAN, two awaiting CRAN release) and has been working on regulatory submission tooling in R for several years. His ORCID is 0000-0001-7864-5819. For this grant he will own the schema specification work in M1, design the validation methodology in M2 with input from data partners, and author the integration vignettes in M3. Relevant skills include R package development, CDISC standards (SDTM, ADaM, Define-XML), Shiny application development, pkgdown and GitHub Actions, and writing JSON Schema specifications.

Aishwarya Amonkar (Techworkslab) is the Co-Lead. She will partner with the Project Lead on technical implementation, validation analysis, and documentation. In M1 she will co-author the JSON Schema spec and write the validator in r4subcore. In M2 she will run the statistical correlation analysis and recalibrate the weights. In M3 she will write and review the pharmaverse integration vignettes. Her skills cover software engineering, technical documentation, R and Shiny development, code review, and quality assurance.

Contributors

Members of the R4SUB GitHub organization (<https://github.com/R4SUB>) who reviewed v0.1.0 and provided early feedback that shaped this proposal. Names will be added before submission as outreach completes.

Consulted

This section will list the people we have spoken to about the proposal before submission, including R Consortium ISC members, members of the R Submissions Working Group, and pharmaverse maintainers. We are following the R Consortium guidance to seek community input before formal submission.

The Problem

What the problem is

A surprising number of R workflows boil down to the same task: collect structured information about whether a set of conditions has been met, score it, aggregate the scores, and make a decision. The terminology varies by community, but the shape of the work does not.

A regulatory team asks whether a submission is ready to file. The R Validation Hub asks whether an R package is qualified for use in a regulated environment ([R Validation Hub 2023](#)). A reproducibility researcher asks whether a study's artifact set meets open science criteria. A data governance team asks whether a dataset is documented well enough to release. In each case the same five steps run in the background. Define what counts as evidence. Define the indicators. Score each indicator. Aggregate the scores along whatever dimensions matter to your domain. Decide.

The R ecosystem has no shared infrastructure for this work. Every community that needs it ends up building it from scratch. Each tool defines its own incompatible schema. Each tool ships its own ad hoc scoring engine. Scores produced in one tool cannot be combined with, audited against, or even compared to scores produced in another. The pattern is the same as before R had DBI for database

connections or `tibble` for tabular data. Different teams solving the same structural problem with mutually incompatible code.

Who it affects

The most immediately affected group is regulatory submission teams. Roughly a thousand organizations submit to the FDA each year, and many more submit to EMA, PMDA, and other authorities (R Consortium R Submissions Working Group 2024). Beyond pharma, the R Validation Hub and the internal validation programs it informs cover any organization using R in a regulated environment, from finance to clinical research. Reproducibility and meta-research groups scoring scientific output completeness face the same workflow. So do data quality and governance teams reviewing whether datasets are ready for downstream analysis or release.

These groups do not need different scoring engines. They need the same scoring engine with different indicator definitions. The fact that each builds its own is wasted effort.

Why it is a problem

Five concrete consequences follow from the lack of shared infrastructure.

Scores are not interoperable. A pharmaceutical company that scores its submissions using one tool cannot meaningfully compare its scores against another organization's, even when both are scoring the same underlying property.

Schemas are reinvented. Every project that touches this kind of work defines its own evidence format with its own vocabulary, which means tooling cannot move between projects.

Decisions are subjective. Spreadsheet driven manual checklists yield inconsistent assessments across reviewers. Two senior reviewers asked to score the same submission will disagree on roughly a third of items in our experience.

Workflows are not reproducible. Ad hoc tools cannot be re-run automatically when evidence updates, so each refresh is a new manual exercise.

Effort is wasted. Tools that solve structurally identical problems are rebuilt across communities that could share infrastructure.

What solving it enables

A standardized R evidence schema and scoring framework gives the R community a common vocabulary for describing evidence and indicators. Scoring becomes reproducible with version controlled weights and decision bands. Cross domain reuse becomes practical, so a tool written for regulatory submissions can be adapted to R package validation by changing the indicator set and the weights, not by rewriting the engine. Decisions move from subjective checklist judgements to quantitative go or no go calls with confidence intervals and sensitivity analysis. The output is a structured audit trail that fits into regulatory submissions, validation packages, peer review, or any other context where a defensible scoring history is required.

Existing work and previous attempts

The closest analogues in the R ecosystem fill different parts of the picture. `pharmaverse` (Nicholls et al. 2024) (admiral, xportr, and others) builds submission datasets and transport files, but it produces artifacts rather than scoring whether they are ready to file. `riskmetric` scores R package quality and risk, which is a single domain solution rather than a general schema. The **R Validation Hub**

(R Validation Hub 2023) provides methodology and policy guidance for validating R in regulated environments, defining an approach but not providing shared scoring infrastructure. **Pinnacle 21** (proprietary) validates CDISC compliance and catches standards violations, but does not give a holistic readiness signal. **R4SUB v0.1.0** is the multi-domain schema and scoring engine, with regulatory submissions as the proof of concept application.

R4SUB is complementary to all of these. It consumes the outputs of pharmaverse and Pinnacle 21 as evidence inputs. It borrows methodology principles from the R Validation Hub white paper. Its risk pillar is conceptually adjacent to riskmetric, with the difference that riskmetric scores R packages and R4SUB's risk pillar scores submission readiness.

Why the R Consortium

This work belongs at the R Consortium rather than inside any single domain or company for three reasons. First, the schema and scoring engine are domain agnostic by design, but each domain (pharma, validation, reproducibility, data quality) has its own funding incentives that would push the project narrow rather than keep it general. R Consortium funding lets the project remain neutral infrastructure that any community can adopt. Second, the R Submissions Working Group has already established that R based regulatory submissions are viable, and the R Validation Hub has established the methodology for assessing R packages in regulated environments. R4SUB is the missing scoring layer that sits across both. The R Consortium is the natural home for infrastructure that bridges working group outputs. Third, the entire ecosystem is open source, MIT licensed, and CRAN distributed. Industry funding tends to produce internal tooling that does not reach CRAN and never benefits the wider R community.

The proposal

Overview

R4SUB is an existing R package ecosystem of eight packages organised around a meta package that provides a single `library(r4sub)` entry point, following the tidyverse model. Six packages are on CRAN today at version 0.1.0. The architecture is functional and tested end to end on its first application, which is regulatory submission readiness for FDA, EMA, and PMDA.

The grant funds three milestones over six months. Milestone 1 takes the evidence schema from an implicit, code defined contract to a formal versioned specification with a CRAN distributed validator. Milestone 2 replaces the current expert opinion based scoring weights with empirically calibrated weights derived from real submission outcome data. Milestone 3 documents how R4SUB integrates with the existing pharmaverse stack and demonstrates that the same infrastructure works in a non pharma domain by walking through R package validation as a worked example.

Milestone	Funding
M1, Schema Standardization	\$3,500
M2, Empirical Validation	\$6,000
M3, Ecosystem Bridges	\$3,000
Total	\$12,500

The R community gets a reusable schema that any package can adopt for indicator based assessment, a scoring framework with empirical rather than heuristic calibration (rare in this space), worked examples of cross domain reuse of the same infrastructure, a lower barrier for any community that needs to assess readiness against indicators, and everything licensed MIT, distributed via CRAN, documented with pkgdown.

Detail

Minimum Viable Product

The MVP already exists. R4SUB v0.1.0 on CRAN runs the full pipeline from evidence ingestion through composite index computation today. The grant funds the work that turns this MVP into infrastructure that other people can confidently build on. The single most valuable deliverable is the JSON Schema specification with a CRAN distributed R validator in Milestone 1. If Milestones 2 and 3 slipped, M1 alone would still produce standardization value that the community can adopt.

Architecture

The eight packages divide along clear responsibilities. `r4sub` is the meta package that loads the others, on the tidyverse model. `r4subcore` holds the evidence schema, parsers, and the validator that this grant will formalize. `r4subtrace` handles ADaM and SDTM derivation chain tracing for the Trace pillar. `r4subscore` runs the composite index calculation and applies weights. `r4subrisk` does FMEA based risk scoring with mitigation tracking. `r4subdata` ships example datasets used by the vignettes and by the test suites. `r4subprofile` holds configurable weight profiles per regulatory authority. The optional packages are `r4subusability`, which adds label and documentation indicators, and `r4subui`, which provides a Shiny dashboard for reviewing scores. Both are scheduled for CRAN release as part of M2.

The flow through the ecosystem looks like this. Domain specific parsers in `r4subcore` (for DefineXML (Clinical Data Interchange Standards Consortium (CDISC) 2021), for Pinnacle 21 reports, and for custom inputs) read source artifacts and emit a schema conformant evidence data frame. Indicator definitions describe how each row of evidence maps to an indicator code and which dimension it belongs to. A profile from `r4subprofile` supplies the weights and the decision bands. `r4subscore` computes per indicator scores, aggregates them into dimension scores, and produces a composite readiness index. `r4subrisk` adds risk priority numbers from FMEA. `r4subtrace` adds traceability coverage. `r4subui` displays everything in a Shiny dashboard. The output is a structured score plus an audit trail of every input that contributed to it.

The schema is deliberately domain agnostic. An indicator in R4SUB has a code, a dimension, a numeric value, a weight, and a metadata block. What is domain specific is the indicator catalogue and the parsers. The current ecosystem ships parsers for regulatory submissions. A different domain can write its own parsers that emit the same schema and reuse the rest of the stack unchanged. This is exactly the pattern that M3 demonstrates with the R package validation worked example.

A worked example

Concretely, suppose a pharmaceutical sponsor runs `define_xml_to_evidence(path)` on their DefineXML file. The parser walks the document and emits an evidence data frame where each row is a single fact, for example: indicator code `Q-DEFINE-001`, dimension `Quality`, value `0.92` (meaning 92% of datasets have non empty `Label` attributes), and a metadata block listing the datasets that were missing labels. A second pass parses the Pinnacle 21 CSV and adds rows for indicator codes `Q-P21-*` covering standards conformance. A third pass walks the ADaM and SDTM metadata and emits rows

for indicator codes T-ADAM-* covering derivation traceability. The combined evidence frame goes to `compute_pillar_scores()`, which applies the indicator catalogue and the active profile, and from there to `compute_sci()`, which produces the composite Submission Confidence Index. The Shiny dashboard in `r4subui` shows the result with drill downs to the originating evidence rows.

For the R package validation worked example in M3, the same engine runs unchanged. A different parser reads `riskmetric` package assessments and emits rows for indicators like RPM-COVTEST-001 (test coverage) and RPM-MAINT-001 (maintainer activity). A different profile applies validation specific weights. The composite index becomes a Package Validation Index instead of a Submission Confidence Index, but the schema, the engine, and the dashboard are identical.

Validation methodology for M2

The empirical validation in M2 takes the current scoring weights, which are based on expert opinion about how much each pillar contributes to submission readiness, and replaces them with weights estimated from real outcomes. The methodology follows three steps. Step one is to recruit two or three pharma partners willing to share anonymized records of past submissions. Each record needs the artifacts that R4SUB consumes (Define-XML, Pinnacle 21 outputs, ADaM and SDTM metadata) and the outcome class (approved, accepted with information requests, or refused to file). Step two is to score each historical submission retrospectively under the v0.1.0 weights, then fit a regression model that predicts the outcome class from the per pillar scores. The regression coefficients become candidate weights for the recalibrated profile. Step three is to validate the new weights against held out cases and publish the methodology so other communities can do the same exercise for their own domains.

If partner data does not materialize in time, the fallback is to use publicly available FDA Refuse to File statistics aggregated by year and therapeutic area, paired with simulated submission evidence calibrated against those statistics. This produces less precise weights than real data but is transparent and reproducible, and the methodology vignette is still publishable.

Assumptions and risks

The proposal rests on four assumptions, each with a mitigation:

Assumption	If false	Mitigation
Two or three partners will share anonymized validation data	M2 cannot use real data	Fall back to public FDA Refuse to File statistics with simulated outcome data
The schema generalizes to non regulatory domains	Cross domain claim weakens	M3 includes a worked R package validation example to prove generalization
CRAN policies remain stable across the project period	Publication delays	Six packages are already on CRAN, so the acceptance pattern is established
pharmaverse maintainers are receptive to integration	Bridge vignettes lack endorsement	Integration is unilateral. Vignettes show how to combine, no upstream changes required

Two further risks worth naming explicitly. There is a risk that the partner outcome data turns out to be small and noisy enough that the empirical weights are not meaningfully better than the expert

weights. This would still produce a publishable methodology but a weaker headline result, and we would adjust the M2 deliverable to focus on the methodology rather than on the specific recalibrated weights. There is also a risk that the R Submissions Working Group views R4SUB as overlapping with their work. We have addressed this by framing R4SUB as a scoring layer that consumes WG outputs rather than as a competitor to them, and by including the WG in the consultation list.

External dependencies

The dependencies are small and stable. The CDISC standards (Define-XML 2.0/2.1, SDTM, ADaM) are public and stable, and they are only used by the domain specific parsers. The CRAN packages we depend on (cli, rlang, tibble, jsonvalidate, ggplot2, shiny) are all mature. The minimum R version is 4.2. There are no proprietary dependencies, no cloud services, and no paid APIs.

Project plan

Start-up phase

The start up phase runs through July 2026 and gets the project ready to start delivering. By July 1 we sign the R Consortium Individual Consultant Agreement. By July 15 we have two or three signed data sharing agreements with our validation partners. By July 22 we have a GitHub project board and a monthly reporting cadence in place. By July 29 the announcement post on the R Consortium blog is live. By July 31 the governance documents (CONTRIBUTING.md, CODE_OF_CONDUCT.md, issue templates) are in place across all eight R4SUB repositories.

All work happens in public GitHub repositories under the R4SUB organization (<https://github.com/R4SUB>) with open issues, pull requests, and discussions. All R4SUB packages are MIT licensed and that does not change. The JSON Schema specification produced in M1 will be released under CC BY 4.0, which is the standard licence for published specifications.

For reporting, we will post monthly progress updates as comments on a dedicated tracking issue in `R4SUB/r4sub-isc-proposal` so the ISC and the community can follow progress without having to ask. The three R Consortium blog posts (announcement in July, mid project in late October, closeout in December) will summarize progress at the natural milestone boundaries.

Technical delivery

Milestone 1, Schema Standardization (August 2026), \$3,500

The current evidence schema lives implicitly in the R4SUB code base. External tools cannot interoperate with it because there is no formal contract to read against. M1 turns this implicit schema into a published, versioned, validatable specification.

By August 8 we publish the JSON Schema specification as `r4subcore/inst/schema/evidence-1.0.json` (draft 2020-12). By August 15 we implement `validate_evidence()` in `r4subcore`, with informative error messages that name the offending row and field. By August 22 we publish the schema reference vignette. By August 26 we publish the parser author guide. By August 31 `r4subcore v0.2.0` is released to CRAN with all of the above.

The success signal for M1 is straightforward. An external developer reading only the published specification and the parser author guide should be able to write a domain specific evidence parser without ever needing to read R4SUB's internal code. We will test this by asking one or two volunteers from outside the project (likely R Validation Hub members or a pharmaverse maintainer) to attempt

to write a minimal parser using only the published spec, and treating any questions they need to ask as gaps in the documentation.

Milestone 2, Empirical Validation (September to October 2026), \$6,000

This is the highest impact milestone and the one that determines whether R4SUB scores can be defensibly used for go or no go decisions. The current weights were chosen by the project lead based on his judgement and on published regulatory guidance. M2 replaces those weights with values estimated from real submission outcome data.

By September 30 the validation dataset is collected from our partners, retained privately per data sharing agreements, and contains twenty or more anonymized cases. By October 10 every case is scored retrospectively under the v0.1.0 weights. By October 20 the statistical correlation analysis (score versus outcome class) is complete and written up as the validation report. By October 25 the recalibrated weights are checked into r4subprofile per regulatory authority. By October 28 the decision band cutpoints are adjusted if the empirical data supports it. By October 31 the validation vignette (`vignette("scoring-validation", package = "r4subscore")`) is published, and r4subscore v0.2.0 and r4subprofile v0.2.0 are released to CRAN.

The success signal is that the scoring model graduates from expert opinion weights to empirically calibrated weights, with the methodology published in enough detail that any community can repeat the exercise for their own indicator set and weights. The mid project R Consortium blog post in late October will describe the methodology and the aggregate findings.

Milestone 3, Ecosystem Bridges (November to December 2026), \$3,000

The final milestone documents how R4SUB integrates with the existing pharmaverse stack and demonstrates that the framework also works outside pharma.

By November 15 the admiral integration vignette (`vignette("admiral-integration")`) shows the ADaM dataset to evidence flow end to end. By November 25 the xportr integration vignette (`vignette("xportr-integration")`) does the same for transport file metadata. By December 10 the cross domain worked example (`vignette("r-package-validation-application")`) walks an R Validation Hub user through applying R4SUB to their indicator catalogue. By December 17 the regulatory traceability matrix is published as a markdown table mapping R4SUB functions to FDA and EMA guidance documents. By December 22 the closeout R Consortium blog post is live.

The success signal for M3 is that an R Validation Hub user, or any other non pharma R community member, can adapt R4SUB to their domain by reading the published vignettes and without help from the project team.

Other aspects

The work will reach the widest possible audience because everything in it is open source from day one. All code is MIT licensed and hosted at <https://github.com/R4SUB>. All packages ship via CRAN, so installation is `install.packages("r4sub")` with no GitHub specific steps required. Every package has its own pkgdown site that updates automatically on commit. Issues and pull requests are public, and we welcome external contributions. The vignettes are linked from the R Consortium blog posts and from each package website.

Our dissemination plan combines the three R Consortium blog posts with conference and social media outreach. The announcement blog post goes live in July 2026 at project kickoff. The mid project blog post on the validation methodology goes live in late October 2026 alongside the M2 deliverables. We will submit an R/Pharma 2026 talk in October and present in November if accepted.

The closeout blog post goes live in late December 2026. Throughout the project we will check in with the ISC at quarterly meetings, post updates on Mastodon (#rstats), LinkedIn, and Bluesky, and cross post each blog post on R-bloggers. We will also offer to present at UseR! 2027 if invited, and at relevant pharmaverse community calls.

Sustainability beyond the grant

The work funded by this grant is designed to be self sustaining once delivered. The schema specification is versioned with explicit deprecation rules so that future changes do not break existing parsers. The validator is a small, well tested function that requires minimal maintenance. The recalibrated weights are checked in to `r4subprofile` so that future updates only need to bump versions. The integration vignettes demonstrate patterns rather than create new code paths, so they remain valid as long as their target packages remain stable.

For ongoing maintenance after December 2026, the project will continue under the existing R4SUB GitHub organization with the project lead as primary maintainer. We are also actively trying to grow the contributor base. The grant funded work explicitly invites external contribution at every step (a published parser author guide, an open governance model, a cross domain worked example), and the goal is to have at least three external contributors making non trivial pull requests by the end of the grant period. The CONTRIBUTING.md and CODE_OF_CONDUCT.md added in the start up phase are part of that story.

Budget and funding plan

The total grant ask is \$12,500. The R Consortium standard payment schedule is 50% on contract signing and 50% on completion. Half of the funds (\$6,250) are released on contract signing and cover M1 (Schema Standardization) and M2 (Empirical Validation) as the work progresses. The remaining \$6,250 is released on final delivery once M3 (Ecosystem Bridges) is complete.

Stage	Amount (USD)
Contract signing	\$6,250
Final delivery	\$6,250
Total	\$12,500

All funds are allocated to labor costs. The estimated effort is approximately 80 hours over six months across the project lead and the co-lead combined, which works out to roughly \$156 per hour fully burdened or about 13 hours per month. The work breakdown is approximately 30% on M1 (specification authoring is detail intensive but bounded), 50% on M2 (the empirical work is the most labor intensive), and 20% on M3 (vignettes and traceability matrix).

There are no indirect costs, hardware purchases, travel, or cloud services in the budget. All the infrastructure we need (GitHub, CRAN, R Consortium blog, R-Universe) is freely available. We have read the R Consortium standard contract for ISC projects (Individual Consultant Agreement, dated 20170622) and accept its terms.

Success

Definition of done

The project is complete when all six of these things are true.

The schema specification is published. JSON Schema v1.0 lives in `r4subcore/inst/schema/`, the `validate_evidence()` function is exported, and two vignettes (the schema reference and the parser author guide) are on CRAN.

The validation report is published. The statistical analysis of R4SUB scores against twenty or more real outcome cases is in a CRAN vignette, with the recalibrated weights checked into `r4subprofile`.

Three integration vignettes are published. One for `admiral`, one for `xportr`, and the non pharma cross domain worked example.

The regulatory traceability matrix is published. A markdown table maps each R4SUB function to the FDA and EMA guidance documents it supports.

All eight packages are on CRAN. `r4subcore`, `r4subscore`, and `r4subprofile` are updated to v0.2.0, and `r4subusability` and `r4subui` are released to CRAN for the first time.

All three R Consortium blog posts are live. The July announcement, the late October mid project post, and the December closeout.

Measuring success

We will track quantitative progress against the following baseline and targets, with monthly check ins on the project tracking issue:

Metric	Baseline (today)	Target (Dec 2026)
CRAN packages published	6	8
Combined CRAN downloads per month	approx 50	250 or more
Unique GitHub contributors across R4SUB org	1	5 or more
External projects referencing the schema	0	1 or more
Validation cases scored	0	20 or more
Test coverage averaged across packages	approx 60%	80% or more
R Consortium blog posts published	0	3
Vignettes across all R4SUB packages	approx 5	approx 12

Quantitative metrics on their own can be misleading, so we will also watch for qualitative markers. A non pharma R community referencing or adopting the schema is a strong signal that the cross domain claim is real. Positive engagement from R Submissions Working Group members or pharmaverse maintainers tells us that the positioning works. Pull requests from contributors who are not the project lead or co-lead tell us that the contribution model is welcoming. Acceptance of an R/Pharma 2026 talk is independent confirmation that the work matters to the community.

If progress against any of these markers is meaningfully off track at the mid project blog post, we will publish that openly along with a recovery plan rather than wait for the closeout.

Future work

The grant funded work is intentionally narrow and self contained. We are deliberately not bundling in everything that R4SUB could eventually become, because the R Consortium funds focused projects

rather than open ended programmes. That said, the infrastructure built in this grant makes a number of extensions tractable that would otherwise require greenfield effort.

The most natural extensions on the regulatory side are additional authority profiles (Health Canada, TGA in Australia, ANVISA in Brazil, MFDS in Korea), CDISC 360 integration to use the linked metadata work for richer traceability, and a hosted Shiny Server deployment for organizations that cannot run R locally. Each of these is straightforward additive work given the schema and engine produced by this grant.

On the cross domain side, the schema and engine are designed to be reused. Concrete extensions we expect community members will pursue include reproducibility scoring (against open science criteria), data quality dashboards (against governance indicator catalogues), and software validation registries beyond R packages. We also expect a deeper integration with riskmetric, where R package risk scores feed into R4SUB's risk pillar as a built in adapter.

On the methodology side, as the validation corpus grows beyond the initial twenty cases, machine learned weight calibration becomes practical. An anonymous benchmarking database across the industry would give individual sponsors a way to compare their scores against an aggregate baseline without exposing proprietary information. Cross language interoperability (Python and SAS bridges) would make R4SUB usable in mixed toolchain organizations.

Each of these could form the basis of a future ISC proposal, a working group effort, or community driven contributions. None of them are in scope for this grant, but each becomes feasible because of it.

Clinical Data Interchange Standards Consortium (CDISC). 2021. *Define-XML 2.1 Specification*. <https://www.cdisc.org/standards/data-exchange/define-xml>.

Nicholls, A. et al. 2024. *Pharmaverse: An Open-Source Ecosystem for Clinical Reporting in R*. R/Pharma Conference Proceedings. <https://pharmaverse.org>.

R Consortium R Submissions Working Group. 2024. *Using R to Submit Research to the FDA: Pilot 4 Successfully Submitted*. <https://r-consortium.org/posts/using-r-to-submit-research-to-the-fda-pilot-4-successfully-submitted/>.

R Validation Hub. 2023. *A Risk-Based Approach for Assessing R Package Accuracy Within a Validated Infrastructure*. <https://www.pharmar.org/white-paper/>.