

SCHARP

Paper OS12
PHUSE US-Connect 2026

Handle with Care: Study-Specific Assessment of R Package Risk

Shannon Grant
Blazej (Moni) Neradilek

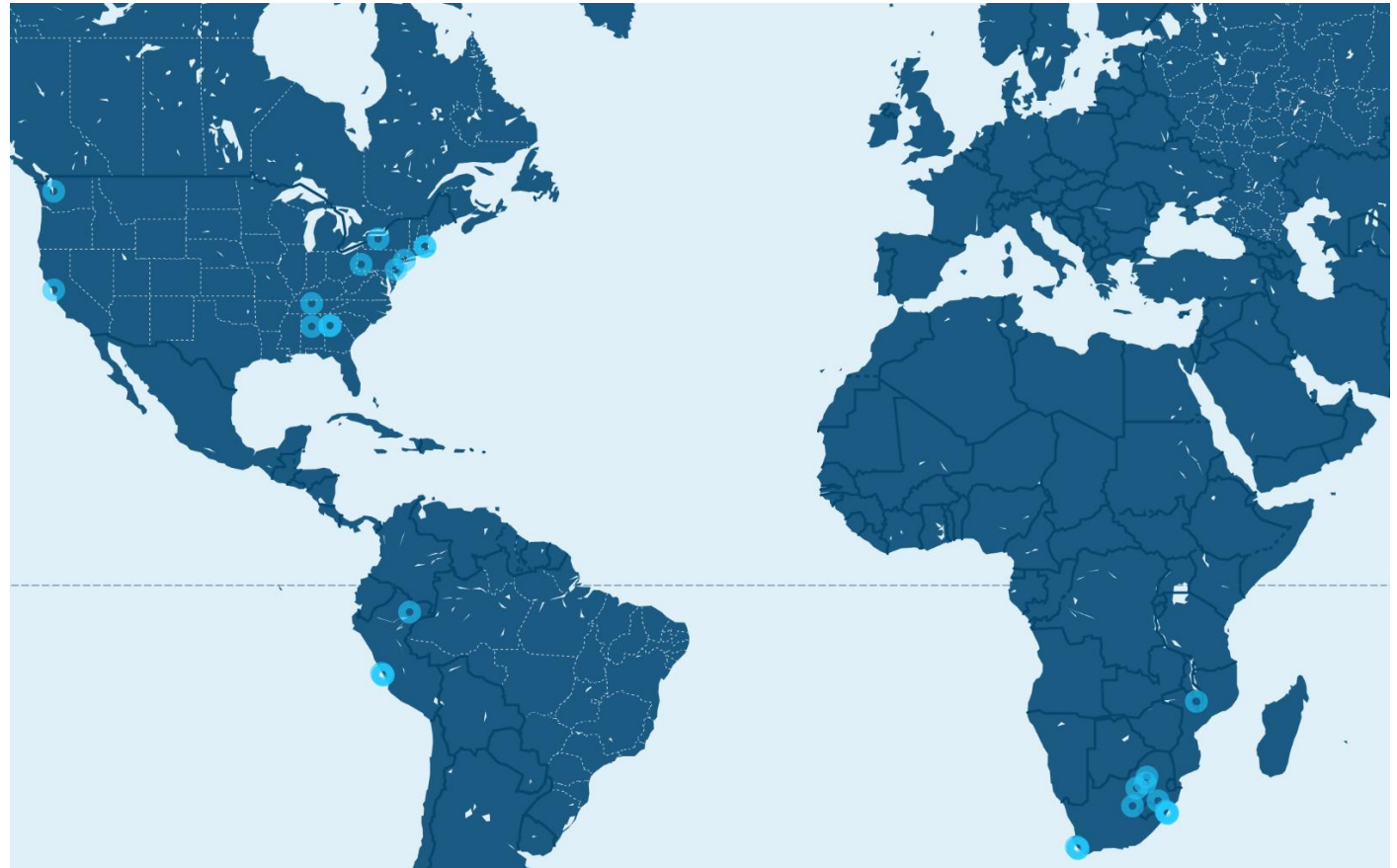
*Statistical Center for HIV/AIDS Research and
Prevention (SCHARP) at Fred Hutch Cancer Center*



- 1 Background and existing guidance
- 2 R validation at SCHARP
- 3 Qualitative risk assessment
- 4 R usage plan
- 5 Conclusions

Who we are

- Nonprofit, quasi-academic medical research institution
- HIV prevention
- HIV Vaccine Trials Network (HVTN) research portfolio is mostly Phase 1
- Past efficacy trials, with more in the future
- Contributed to submissions...in SAS
- HIV vaccine statistical community strongly prefers R



HVTN site map from <https://www.hvtn.org/participate/study-clinics.html>

R validation challenges

- Software should be “reliable”, and computer systems validation should “should be based on a risk assessment that considers the intended use of the system” (ICH E9, ICH E6)
- Validation includes Installation, Operational, and Performance Qualification testing (IQ, OQ, and PQ)
- **In PQ, how can we prove that R packages are reliable and accurate?**
 - Base R and recommended packages
 - Community-developed packages

Guidance from R Validation Hub

R Validation Hub white paper:

- System validation should include accuracy, reproducibility, and traceability
- Risk assessment framework for R packages should include four dimensions:



Purpose



SDLC



Community
usage



Testing

Implementation:

- `riskmetric` package to collect package characteristics in these 4 dimensions
- Case studies collected by R Validation Hub incorporate these characteristics, but vary w.r.t. actions taken and how risk is assigned (algorithm vs expert judgement)

Appetite for risk

- How far does a team go to formally assess R packages?
- Balance between resources and concern about errors/consequences
 - Most people want their work to be correct
 - Not everyone has pharma-level funding or impact
- A range of risk is acceptable at most organizations
 - High-impact/regulated areas have formal quality procedures
 - Exploratory/research projects are more flexible
- SCHARP has moderate risk appetite
 - NIH-funded clinical trials, formal Quality system, but mostly research-track with a nonprofit budget



R validation at SCHARP: Our environment

- HIV vaccine statistical community strongly prefers R, but staff use a mix of SAS and R
- SCHARP Statistical Computing Environment (SCE) on a Unix server is managed by a small IT team



R validation at SCHARP: practical PQ solution

Two steps to documenting confidence in chosen R packages:

1. Streamlined **R PQ** for validated packages
 - Library used across the organization
 - Traditional PQ not practical
 - Single PQ test case: no errors in R CMD checks*
 - PQ occurs at installation in validated library
2. **R Usage Plan** is study-specific
 - Statistical study leads understand their R needs
 - Risk assessment & mitigation occurs during study planning

**Manual review and judgement applied to decide if packages not passing checks can be included in validated library*

R usage plan: Qualitative risk assessment

Selected package characteristics:



- Number of reverse dependencies



- Number of annual downloads



- Bug report URL exists



- Help files present



- Trusted source (base R, recommended, Tidyverse)



- Passing R CMD checks



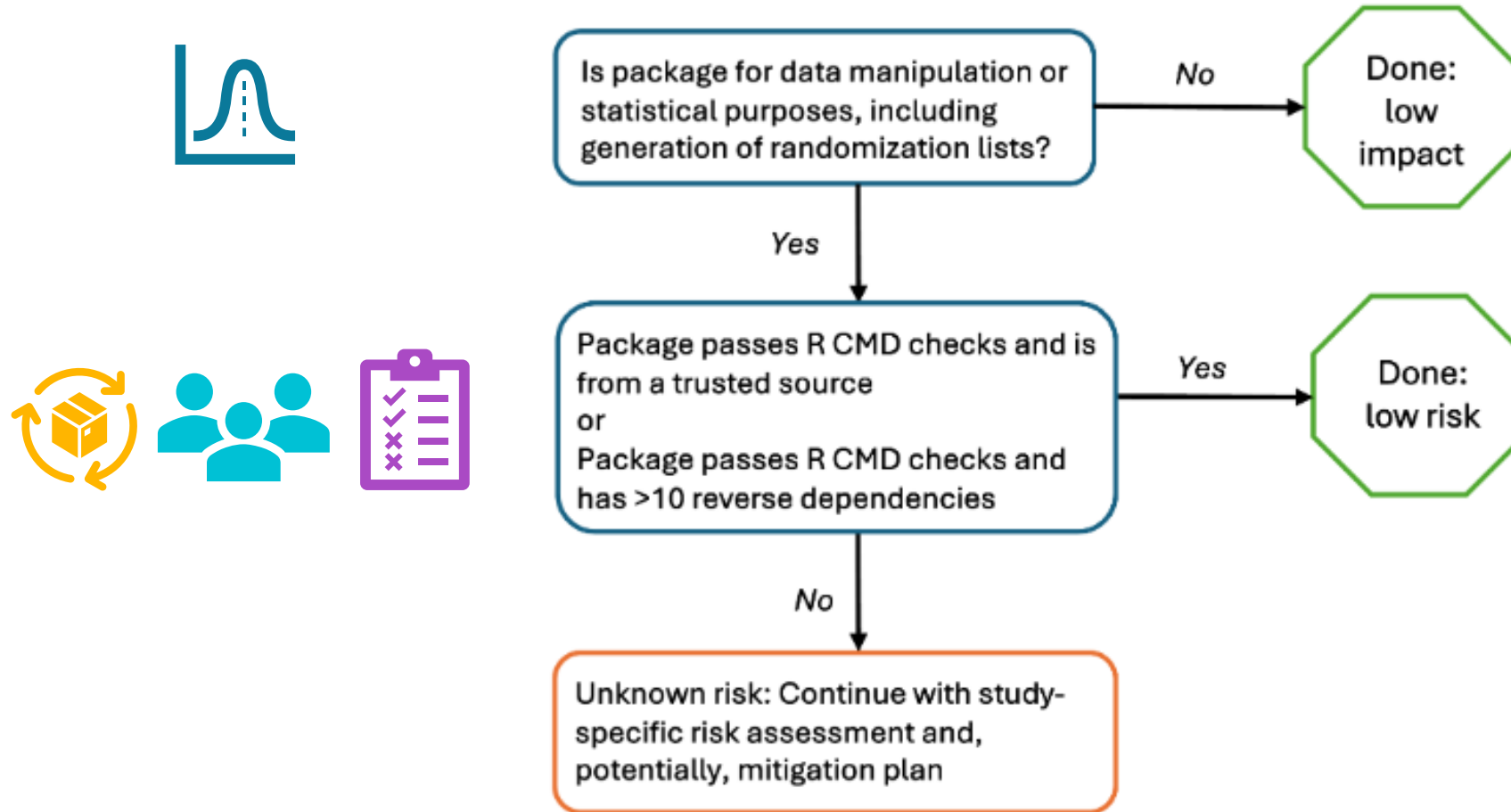
- Purpose

R usage plan: Overview

- R usage plan content
 - Initial risk assessment (automated)
 - Qualitative study-specific risk assessment
 - If needed, study-specific risk mitigation planning
 - Reproducibility plan (*out of scope for today*)
- Scope: relevant functions for included packages

Required	Optional	Excluded
• Primary endpoints • Generation of randomization lists	• Trial design • Secondary endpoints • Dependencies	• Verification, data cleaning, other QC

R usage plan: initial risk assessment



R usage plan: Study-specific risk assessment

No required criteria, but we recommend thinking about the following.



- Likelihood of encountering problems
 - Developmental, new, or little-used packages are riskier



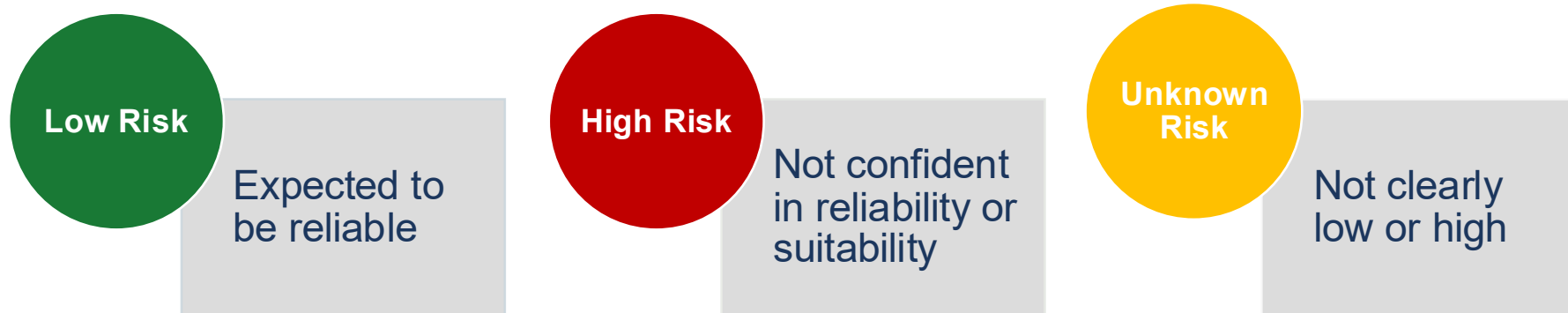
- Ease of troubleshooting
 - Documentation and SDLC are reassuring
 - Lack of clear SDLC may be balanced by strong history and community use
- Suitability of the package for the study operating characteristics
 - Statistical methods for ties, censoring, event rates...



- Unit test coverage
 - If available and effort is justified

R usage plan: Study-specific risk conclusion

- Statistical team considers their **risk appetite** and the **balance of package characteristics** in study context when deciding on final risk level
- High and unknown risk go on to risk mitigation planning



R usage plan: Risk mitigation

- **Consider:**

- Package characteristics
- Impact of errors on the study
- Study team's risk appetite
- Resource availability

- **Potential mitigation steps:**

- Compare to known results
- Identify or use an alternative
- Custom tests
- Do nothing

Example: Initial risk assessment (and beyond)

Table 3: Initial risk assessment for external packages. Sorted by initial impact/risk, purpose and package name. *(continued)*

Package			Deliverable and objectives		Low impact criteria	Low risk criteria			Potential study-specific risk assessment characteristics, if applicable			Initial impact/risk assessment
Name	Version	Validated	Deliverable(s)	Objective(s)	Purpose	CRAN/Bioc check error	Source	Reverse dependencies	Annual downloads (in 1000s)	Bug report URL exists	Help files present (%)	
purrr	1.0.2	Yes	PT report, pdata	P2, S1	Data manip.	No	Tidyverse	2099	11492	Yes	100	Low risk
survival	3.8.3	Yes	PT report	P2, S1	Statistical	No	Recommended	1086	1549	No	100	Low risk
Exact	3.3	No	PT report	P2, S1	Statistical	No	Other	5	525	No	100	Unknown risk
MIMOSA	1.29.1	Yes	pdata	S1	Statistical	No	Other	0	<1	No	67	Unknown risk
PropCIs	0.3.0	Yes	PT report	P2, S1	Statistical	No	Other	3	14	Yes	100	Unknown risk
randomizeR	3.0.2	Yes	Randomization list	NA	Statistical	No	Other	1	6	No	78	Unknown risk

P2 means Primary objective 2, S1 Secondary objective 1. *Other indicates packages with other non-statistical, non-data manipulation purposes.

Example: Study-specific risk assessment and mitigation plan

4. `randomizeR`

- **Strengths:** This is a mature external package, first released on CRAN in 2015, and described in the Journal of Statistical Software (Uschner *et al*, 2018). The currently available version 3.0.2, published in 2023, is included in the validated R installation at SCHARP.
- **Weaknesses:** This package has a small number of reverse dependencies (1) and downloads (6255 in the last year). Help files are not complete, covering 78% of exported objects. However, closer inspection of the package using function `getNamespaceExports` reveals that 85 exported objects exist, but only 66 are user-facing, and the rest are internal/hidden objects. After excluding these internal/hidden objects, help file coverage is 100%.
- **Conclusion: Unknown risk.** While this is a mature package with accepted peer-reviewed methodology, it is not that widely used, and does not have complete documentation. Problems could arise which are difficult to resolve.
- **Mitigation:** No mitigation is needed. The randomization list will be checked according to SCHARP WI-0005 Randomization, which includes verifying that the number of treatment groups, strata, blocks, block size(s), and number of treatments within a block are as specified. If any issues with `randomizeR` arise which cannot be resolved, SAS PROC PLAN can be used instead.

R usage plan: Efficiencies and team learning

- R usage plans are governed by a new SOP with many examples
- R usage plan template makes it easy
 - Rmarkdown report, code, and input templates
 - Package characteristics from R PQ, or use similar code for non-validated packages
 - Study-specific examples for users to edit
 - Managed and distributed internally using GitHub.com

Conclusions

- New process! We will evaluate and modify over time
- Template, instructions, and examples make it easier
- Qualitative approach risk assessment
 - Requires thinking about package-specific risks
 - Serves goals of learning and preparing for higher scrutiny



References

- International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. *Good Statistical Principles for Clinical Trials* (ICH E9, 1998). https://database.ich.org/sites/default/files/E9_Guideline.pdf, accessed 2 January 2026.
- International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. *Guideline for Good Clinical Practice* (ICH E6(R3), 2025). https://database.ich.org/sites/default/files/ICH_E6%28R3%29_Step4_FinalGuideline_2025_0106.pdf, accessed 2 January 2026.
- The R Foundation for Statistical Computing (2021). *R: Regulatory Compliance and Validation Issues A Guidance Document for the Use of R in Regulated Clinical Trial Environments*. <https://www.r-project.org/doc/R-FDA.pdf>, accessed 2 February 2025.
- Gillespie CS. *Can I trust that package?* (posit::conf 2025). <https://www.youtube.com/watch?v=g6QU16jOt0g>
- **R Validation Hub, Nicholls A, Bargo PR, Sims J. *A Risk-based Approach for Assessing R package Accuracy within a Validated Infrastructure* (2020). <https://pharmar.org/white-paper/>, accessed 2 February 2025.**
- R Validation Hub, Kelkhoff D, Gotti M, Miller E, K K, Zhang Y, Milliman E, Manitz J (2026). *riskmetric: Risk Metrics to Evaluating R Packages*. R package version 0.2.6, <https://CRAN.R-project.org/package=riskmetric>.
- Case studies from collection at <https://pharmar.org/categories/case-studies/>.
 - Roche R Package Validation Team. *Automated R package validation*. https://github.com/pharmaR/case_studies/blob/fa5e4c82098266fe7516adaa8671a64c9796a0e3/Roche_Rpkg_validation_2022.pdf
 - Stachniuk A, Wallet P, Fidler M, Ochsenbein D, Schmidt S, Illis W. Case study: *R Package Risk Assessment at Novartis* (2022). https://github.com/pharmaR/case_studies/blob/fa5e4c82098266fe7516adaa8671a64c9796a0e3/Novartis_Rpkg_validation_20220509.pdf
 - Palukuru UP, Bernecki P, Liao J, Zhang Y. *External R Package Qualification Implementation at Merck: Case study using GGally* (2022). https://github.com/pharmaR/case_studies/blob/fa5e4c82098266fe7516adaa8671a64c9796a0e3/External%20R%20Package%20Qualification%20Implementation%20at%20Merck.pdf
 - Manitz J, Pinkert S, Gregory M, Beckers F. *Case Study: Risk Assessment of R Packages and Merck KGaA/EMD Serono* (2022). https://github.com/pharmaR/case_studies/blob/fa5e4c82098266fe7516adaa8671a64c9796a0e3/Merck_KGaA_Rpkg_validation_220202.pdf
- Wickham H, Bryan J. *R Packages* edition 2, Appendix A. <https://r-pkgs.org/R-CMD-check.html>, accessed 20 January 2026.
- Photos: [Yuliya Strizhkina \(Cartier\) UA](#), [Tai Ngo](#), and [Erda Estremera](#) on [Unsplash](#)

Acknowledgements

- Thanks to our colleagues from the SCHARP R PQ team, whose work and insight lay the foundation for this paper: Radhika Etikala, Valeria Duran, Dave Slager, Xuehan (Emily) Zhang, and Amy Harper.
- Thanks also to Jiani Hu for programming the first draft of the R usage plan template.

SCHARP

Questions?

Thank you!

Shannon Grant (sgrant@fredhutch.org)

Blazej (Moni) Neradilek (bneradil@fredhutch.org)

