

R “Validation” Hub Meeting

19 Nov 2019

Attendees:

Doug Kelkhoff (Roche)– Meeting Chair
Lyn Taylor (Phastar) – Meeting Secretary

Bob Engle (Biogen)
Evelyn Du (Teva)
Ian Wallace (Cel Gene)
Joe Rickert (R Studio)
Juliane Manitz (Merck Serono)
Keaven Anderson (Merck)
Matthias Trampisch (Boehringer-Ingelheim)
Melvin Munsaka (Abbvie)
Nash Delcamp (Cognigen)
Niccolo Bassani (Quanticate)
Noam Ross (EcoHealth Alliances)
Patric Stracke (Sanofi)
Paul Schuette (FDA)
Paulo Bargo (J&J)
Pieter-Jan Stiers (GLPG)
Raj Malathker (J&J)
Rebecca Krouse (RHO Inc.)
Reinhold Koch (Roche)
Rose Hart (Bresmed)
Yilong Zhang (Merck)

Agenda

- R Consortium funding update
- White paper sharing and input request (Paulo)
- Call for metric discussion contributors, possible introduction of a new workstream (Doug)
- Additional updates from streams as appropriate

R “Validation” Hub Meeting

19 Nov 2019

Previous Action Items

Action Item	Assigned team member(s)	Deadline	Status
Communication, dissemination of information about the project to the wider group <ul style="list-style-type: none">- Submit application for funding to develop a shiny app.	Andy Nicholls/ Lyn Taylor/ Joe Rickert	Sept 2019	Complete. Awarded
Review of the website content: Direct review back to parker via trello or slack. Possible future action to get the R consortium website to point to our minutes on our website. Update to include roadmap and progress. <ul style="list-style-type: none">- Suggestions for change of name – Please!	Kieran Martin, Min Lee, Juliane Manitz, John Simms, Keaven Andersen, Parker Sims	Oct 2019	Ongoing
Risk assessment workstream – Continue work on rickmetrics package – with input from the wider team	Yilong Zhang , Rebecca Krouse, Doug Kelkhoff, Matthias Trampisch, Eric Nantz	Oct 2019	Ongoing
Requirements/tests workstream set objectives + milestones - ACTION: Tilo share tests via GitHub	Keaven Anderson , Nate Mockler, Tilo Blenk .	Oct 2019	Ongoing
Validation white paper workstream – Target publishing of white paper by 18 th Oct	Andy Nicholls, Paulo Bargo, Lyn Taylor, John Sims	Oct 2019.	Ongoing

Discussion

- **R Consortium funding update**

\$16,000 awarded by the R consortium for our group to spend on developing a website that will house the Risk metrics information online.

There will be a contract put together and we need to organize transfer of the funding. We need to meet the milestones that were put together in the proposal to receive the funding.

We will need to finalize a version of the Risk metrics format and then contract someone to develop a website to house that online. We need to come up with our own metrics that we want to see on there and have it integrated with the structure we have. The idea would be to build a database that can be accessed by anyone.

We need to put together a description of our project for the R consortium website. Which could come direct from the proposal. Could put the proposal itself online or a description onto the website to announce what we are doing.

ACTION: Joe & Andy: Communicate award & initiate team to put detail on the website

ACTION: Joe & Andy: Put together more targeted milestones and next steps.

R “Validation” Hub Meeting

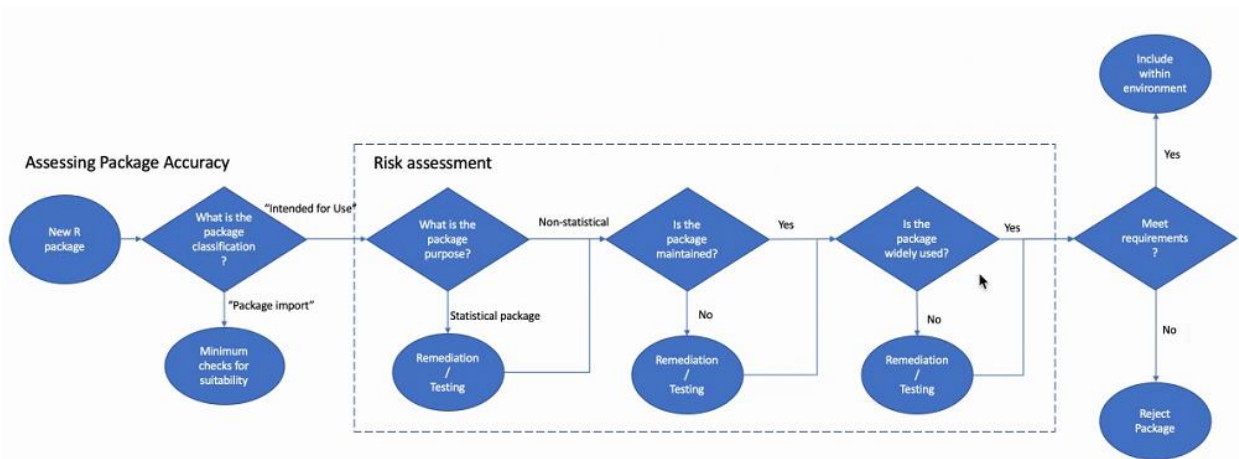
19 Nov 2019

- **White paper sharing and input request (Paulo)**

The team spent many months putting together “The R Validation white paper”, which details similar content to the website and directs people to the website to allow the paper to be relevant even when future updates are completed.

Mostly Paulo, John Sims & Andy worked on it.

In the meeting Paulo went through the white paper and invited comments from the team. See the white paper for more detail. Prior to finalization a diagram detailing the flow process, will be added. Draft example shown below:



Matthias Trampusch asked what the plan was for imports of other packages. i.e. Suppose dplyr is getting input by GGLOT2, do we have to put both dplyr and GGLOT2 through the process? If for example, a package imports 80 other packages and one of them doesn't pass the process flow, does it mean you have to reject all the higher level packages?

Joe responded explaining that the flow chart doesn't preclude people modifying it, it's not to be enforced, but a guidance to show the level of risk assessment which is minimal. If you have some low level package that has a high level of visibility, it's high chance that the specific part which is called has a track record of it being used in that function and scrutinized as part of that package. This is different to using the package on its own under full function.

Matthias responded that he struggles to understand the use of different rules for risk, based on intent to use or Import a package.

Paulo added that you could assess the 80 packages working together as a whole – you do not have to assess all 80 packages but you may want to. The paper does not advise against full assessment of all packages or insist on it, it's just a risk assessment. If a package being imported is considered to have risk, then you can do a risk assessment on it. In general, packages imported by other packages don't have to go through such rigorous testing, however, if the package import is being used for statistical inference or key part of the analysis, then we may decide to assess the risk of it in a more thorough way. The guide is more a minimum that you should do, but not limited to only this. You should do assessment and testing based on your needs.

R “Validation” Hub Meeting

19 Nov 2019

The alternative is that you have to run all the packages through the same scrutiny which would invest a lot of resources and perhaps not be feasible.

Doug questioned the trusted Sources section; he has an alternative view on a “trusted “ source. How did the team come to agreement on this list.

Paulo responded that it’s an example, and just because a package is done by a certain trusted group doesn’t guarantee it’s flawless. However, you find bugs over time so whether you trust it or not, falls into the same assessment of risk. How much additional testing do you feel that you have to do. Something like tidyverse you would test in your environment, however you shouldn’t need to do thorough tests that you would have to do to a package doing statistical analysis.

Doug informed the group that in developing the riskmetrics package, they aim to make it as automated as possible. He would like to see the recommendation viewing everything in a holistic way. If the riskmetrics package can be automated, then maybe we could run it against all the imports as well.

Paulo responded that although that could be done by bigger companies, it still may not be possible or desirable for smaller companies. It’s the risk that the person is willing to take, and this cannot be enforced by this group. It’s an individual assessment because it’s them that has to respond to the authorities. Whilst the paper could have been more strict, it was felt better to be written as minimum checks for suitability. In future, we’ll have to be more fluid in the webpage and perhaps do recommendations that are more tight. However, this wont go into the paper as its more of a general concept.

Matthias asked if an import package does the stats testing, then perhaps that would need to be risk tested if it’s the main working package. He requested to add that as add a caveat into the white paper.

Paulo was asked if the team will be putting this version of the paper up for wider review. Lyn suggested that rather than delay release, it would be better to release a version and then invite comments and revise in future. There was interest in the group to review & comment prior to publishing. Doug and Joe agreed to take it back to the executive committee.

ACTION: Joe/Doug to discuss with Andy & come up with a plan for release and review / revision.

- **Call for metric discussion contributors, possible introduction of a new workstream (Doug)**

Validation risk assessment page:

- The latest version is on github – if you go on issues & sort by label, you can look through issues.
- Can the team try to look through these and assess what your own company internal validation is like and check that we have covered here what you would do. If not then please go to “Suggest a new metric” which will add an issue and will automatically get tagged as a metric issue.

Question to team: If this a reasonable way to get feedback from our organizations as to what metrics we need to incorporate.

R “Validation” Hub Meeting

19 Nov 2019

Question to the team: does anyone know how their company assesses risk and as such what metric we should be assessing. If not, then we should get those people involved from our companies. We should have perhaps the regulatory department involved to give us that kind of feedback on the metric we are collecting.

Given the lack of initial response on these items, Joe speculated that the people on the call, likely don't get involved with the setting of the testing & metrics and maybe we need to put out a request to say enable us to talk to the right people in the organizations.

It was also speculated that people may be afraid to talk openly about their organization.

Yilong commented that in the end we will have recommendation /guidance to help us move forward so it might be easier to release and then get feedback. In general within Merck, they rely on guidance from community, and identify risk based on consensus from the community. Currently Merck are very conservative, the only packages of low risk are considered to be the R foundation packages. Tidyverse is moderate risk and not to be used for regulatory deliveries.

Doug asked if it is the case that other companies are not using packages that are from the community. Matthias responded that they check the basic metrics from cran and then make a decision on each package. If a function of a package is used, they require a full validation such as comparison to another package or SAS proc step. He added that this can be cumbersome & takes considerable time. All of the metrics are new to them. Not all probably important or relevant to the outcome of the assessment, but something like this has not been done at a lot of companies. Decision making (Use / not use) is done by validation and they have to check that the output of the calculation is correct. They could not simply do a risk assessment using metrics and then use the package.

Doug asked if they implemented unit tests for the metrics, would that be able to be used. Matthias wasn't sure if you can apply a unit test for very technical stats packages. He suggested that you must validate the actual result.

Joe concluded that this is ongoing work, and it's possible to achieve a large amount of unit testing on some packages especially if some packages are adopted by the community. In these cases, then the consortium could start the process of possibly awarding grants to work further on the testing and documentation of that.

Doug agreed that this avenue would be good to discuss in future.

Doug informed that group that if you are interested in taking part of that discussion visit the github repo.

R “Validation” Hub Meeting

19 Nov 2019

pharmaR / riskmetric

Unwatch 8 Unstar 14 Fork 5

Code Issues 31 Pull requests 1 Actions Projects 1 Wiki Security Insights

Pinned issues

- Metrics Wish List #8 opened on Mar 22 by dgkf Open 6
- package layout #2 opened on Mar 7 by kieranjmartin Open 23

Filters is:open is:issue label:"Metric Proposal" Labels 13 Milestones 4 New issue

Clear current search query, filters, and sorts

17 Open 1 Closed Author Labels Projects Milestones Assignee Sort

- Metric for percentage of issues open/closed/with response on bug tracker Difficulty: Medium Metric Proposal #56 opened on Sep 24 by dgkf
- Metric for presence of a maintainer Difficulty: Low Metric Proposal #55 opened on Sep 23 by dgkf
- Metric to assess number of active contributors Difficulty: High Metric Proposal #54 opened on Sep 23 by dgkf
- Metric for gauging suitability of a software license Difficulty: Low Metric Proposal #53 opened on Sep 23 by dgkf 2019.11 3
- Metrics for history and number of package downloads Difficulty: Medium Metric Proposal #52 opened on Sep 23 by dgkf 2019.11 8
- Metric for assessing source code repository community enthusiasm Difficulty: Medium Metric Proposal #51 opened on Sep 23 by dgkf

Additional updates from streams as appropriate

- No time for further discussion on this item.
-
-

Actions

Action Item	Assigned team member(s)	Deadline	Status
Review of the website content: Direct review back to parker via trello or slack. Possible future action to get the R consortium website to point to our minutes on our website. Update to include roadmap and progress. Suggestions for change of name – Please!	Kieran Martin, Min Lee, Juliane Manitz , John Sims, Keaven Andersen, Parker Sims	NA	Ongoing

R “Validation” Hub Meeting

19 Nov 2019

Risk assessment workstream – Continue work on rickmetrics package – with input from the wider team	Yilong Zhang , Rebecca Krouse, Doug Kelkhoff, Matthias Trampisch, Eric Nantz	NA	Ongoing
Requirements/tests workstream set objectives + milestones - ACTION: Tilo share tests via GitHub	Keaven Anderson , Nate Mockler, Tilo Blenk.	NA	Ongoing
Validation white paper workstream – Target publishing of white paper by 18 th Oct	Andy Nicholls, Paulo Bargo, Lyn Taylor, John Sims	NA	Ongoing