R "Validation" Hub Meeting

21 Jan 2020

Attendees:

Andy Nicholls (GSK) – Meeting Chair
Lyn Taylor (Phastar) – Meeting Secretary

Carl Walker (Corteva Agriscience)
Chris Toffis (SQN Clinical)
Doug Kelkhoff (Roche)
Ed Lauzier (Merck EMD Serono)
Evelyn Du (TEVA)
Greg Cicconetti (Abbvie)
Iain Wallace (Cel gene)
John Green (Eco Stats)
Joseph Rickert (R Studio)
Juliane Manitz (Merck Serono)
Keaven Anderson (Merck)
Lin Taft (GSK)
Mark Padgham (Eco Health Alliance)
Melvin Munsaka (Abbvie)
Michael Cartwright (parexel)
Nate Mockler (Biogen)
Niccolo Bassani (Quanticate)
Noam Ross (Eco Health Alliance)
Patric Stracke (Sanofi)
Patrice Kiener (Inmodelia)
Paul Schuette (FDA)
Paulo Bargo (J&J)
Phil Bowsher (R Studio)
Phillip Clarke (RCPE)
Pieter-Jan Stiers (GLPG)
Rebecca Krouse (RHO Inc.)
Søren Klim (Novo Nordisk)
Tomas Drgon (FDA)
Xiaoyi Sopko (Corteva Agriscience)

R "Validation" Hub Meeting

21 Jan 2020

Agenda

- New Logo has now been released and will be appearing on the website shortly.

- Updates on workstreams (Workstream leads)
- White Paper Dissemination (Andy)
- ROpenSci and statistical software (Noam and Mark)
- AOB

Previous Action Items

Updates

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
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 18th Oct	Andy Nicholls, Paulo Bargo, Lyn Taylor, John Sims	NA	Ongoing

R "Validation" Hub Meeting

21 Jan 2020

Updates on the workstreams:

Website: Juliane Manitz

The website is being updated following some pull requests. They still need to add the funding summary, and it's still to be confirmed if they should link to the proposal itself from the website or just summarize it on the website? They will also publish the white paper, and update logo on the site.

Given we are now nearly at 100 members across 60 companies, we are keen to have the website represent that we are an official group with wide representation

Risk assessment workstream: Doug Kelkhoff

The onboarding material has been cleaned up and process defined for contribution. We are trying to come up with the list of metrics we plan to include.

Requirements workstream: Keaven Anderson

Most work to date was proposed by Tilo based on work at GSK to do a qualification of an installation. The concept would be that once you are happy with the package risk, how would you then quality the install is as you expect. The team hope to be able to share this at the next meeting.

White paper: Andy
The white paper in its raw form is available on github. It talks about how you might mitigate the risk through testing, if a package is considered high risk. The paper is not specific on the type of testing, but it's generally considered that for stats analysis packages testing would be needed.

Changes since the last version which was shared by Paulo at this meeting, are clarification that the white paper is our current thinking and it might evolve.

Next Steps: The white paper can now be put onto the website, then Andy will reach out to the team via email to inform everyone that it's released, and we hope to have a corresponding article in PSI News (SPIN) with a summary on the white paper. This higher level summary might be useful for you to share with colleagues as it will be a abbreviated summary of the white paper.

ACTION: ALL: Share the white paper and thinking within your companies and any feedback (agreement or disagreement) is useful for the team to know.

R "Validation" Hub Meeting

21 Jan 2020

ROpenSci and statistical software (Noam Ross and Mark Padgham)

Noam Ross introduced himself as a computation biologist working on the ROpenSci project. This is an organization composed of largely academic researches and open source developers in R and statistics. It's purpose is to support researchers to produce open and reproducible research. The team have support through open source communities and documentation. For past 5 years, the project has operated a pear review of packages, similar to review of open source code and manuscript review. They've handled close to 250 packages now, developed standards and testing tools to support reviewers and authors who develop them. They give a "stamp of approval" or badge and have partnership with some journals who also stamp approval of the package. The reviewers look at line by line code, package documentation and at testing suites. The hope is that it drives better testing and development of packages and also leads to better packages shared in the community.

In the past they have avoided packages that create stats analysis, focusing more on non-statistical packages. However, this has changed late last year where they are expanding peer review system to include R packages that have implementations of statistical algorithms and produce statistical outputs. They are re-building their peer review system to cope with this. Re-usable standards, testing tools & review journal.

Mark Padgham explained to the group that he joined the project last year and is in charge of leading and developing the new part of the ROpenSci project that will look at the stats packages. The project in the past has developed organically so this is an opportunity to go back and look at the structure. There is a lot of overlap with R Validation Hub's activities, since they will also be developing a suite of tools for testing and validation. They are confident that the riskmetrics package will be something they use and hence would be happy to collaborate with us on this – especially because it's been developed in such an open and flexible way. Will work closely with Doug/Yilong on agreeing set of metrics.

ROpenSci will be using a badging system with an assessment of risk but what they are doing is more general assessing not just the risk, but other properties and metrics of output – a "been reviewed by rOpenSCi badge". New system will be a peer review (like a journal review) which after the badge awarded it can be cited. As well as the initial badge, they expect an ongoing micro review to ensure it remains up with current standards throughout its lifecycle, they envisage a combination of peer review and micro review to ensure the software is kept up to date with ever changing standards.

ROpenSci are also working on the equivalent to our white paper, it's available on github in draft, but they hope in a couple of weeks, to be able to share a link to that document. The white paper contains standards and operational context of the project. There is a steering committee of 8 people now set up and they will have their 1st meeting next week. The hope is that they will have a great deal of input into the white paper. Given the assessment of software is broad and vague the steering committee will try to agree on a set of metrics they will measure. One divergence between the R Validation Hub and ROpenSci is that RopenSci may look at R packages, but also may broaden to other software using R code or other code such as python.

Timelines for stats package project is 2 years and it started at the end of 2019. This year will begin with the release of the white paper in 2 weeks time, and they hope by the end of 2020 to have a general vision of structure of the project – what software will be accepted and what packages will be looked at. During the second half of 2020, they expect to create a set of testing support tools – that users can confirm or validate. The goal is to do a peer review of a stats package at the start of 2021 and start accepting packages (i.e. taking 1st packages through the process) in 2021.

An ISC working group will be set up in parallel to the R Validation Hub, and they will meet every other month to work on the standards and test suites to create an online forum and calls, to share across various teams. The hope of this ISC group is to coordinate and avoid duplications across the different teams.

Joe Rickert pointed out that even with the standards to test the stats packages, you'll have to rely on experts in particular areas to do the testing. How will this be able to scaled up and how many packages will you be able to review in a time period? How would you prioritize packages? Noam responded that they will draw on the packages they've done already, driven by the authors who submit them to be assessed. Initially they got around 10-15 per year, but now they are up to 30 per year so they've had to scale up on the team. They have volunteer editors who use their networks to find volunteer reviewers. There is also automation that they can try to implement in the process, such as email reminders, identifying reviewers based on background and expertise. It is very much a process that required human expertise/intervention. The other side of tooling they want to support is to have a set of tools for authors that allow them to do annotation of their code to show where and how it meets the standards. This may help efficiency and scale. It scales with the number of people that want to get involved.

Doug Kelkhoff raised concerns with whether it would be easy to spoof a badge and have the team thought of any way to avoid this? However, as part of ROpenSci plans, each badge links through to the review that occurred and diagnostics. Continuous integration shows that all packages that have gone through peer review to show any that did fail in any area and why, especially useful if due to a standards change. People can see which version passes and when it does, as well as any time it didn't pass. Therefore, they aren't concerned about malicious actions, more concern about keeping things up to date. The links back to review process gives it a lot of audit tracking, viewable beyond the individual badge assigned.

Doug also asked how do they demonstrate the precision requirement of the stats tests. One example that will be incorporated is a reference to the NIST standards (national institute of tests), they provide test data and identify their standards to a specific precision, so you can apply the tests at the precision given and use that as evidence of accuracy. Any software that uses this can reference this precision.

Andy Nicholls agreed that there are parallels between the groups especially with regards to the approach for stats packages, and ROpenSci white paper will help us in the hub. The R validation Hub's longer term roadmap is to create a package repository suitable for pharma, where a package is accepted if they've met a set of standards. Therefore, having ROpenSci putting together a badge and standards, it will be really helpful.

Per the white paper, Andy raised the concept of a "Trusted resource" for industry, similar to how we consider SAS. Perhaps the badge of ROpenSci could be considered a trusted resource, and based on the award of a badge from ROpenSci means it's trusted for use in pharma and regulatory research.

ROpenSci has other projects that may be useful for repositories which we could collaborate with such as hosting the repository under a home space collection on github.

It needs to be agreed on how the tools and metrics contribute back to the open source package. Doug & Noam will work together on this. Noam will have their draft framework available soon which Doug can review and give feedback on. Independent of this, Noam is looking for more input from different model systems, including different audit processes to see different examples of how similar audits are or are not. He'd like to draw lessons not from only open source world but from Pharma audit processes.

ACTON: All (Doug/Noam) to collect examples of audit processes in pharma to demonstrate the non open source area to see if this can be considered by R open Source.

AOB:
Carl Walker and Andy spoke last week about a potential overlap with agri-science. A lot of what is being done is not unique to pharma, there is overlap with finance, insurance, and agri-science which are all regulated industries. Carl's company creates biocrops, for pharma use so it's heavily regulated. Although there will be differences in the packages they use, there is sure to be overlap in the validation approach.

In agri-science, they want the regulatory bodies to be using the same packages and hence to centrally validate the packages, rather than individually company validating them and selecting which to use.

Actions

Action Item	**Assigned team member(s)**	**Deadline**	**Status**
Website content updates	Kieran Martin, Min Lee, Juliane Manitz, John Simms, Keaven Andersen, Parker Sims	NA	Ongoing
Risk assessment workstream – rickmetrics package	**Yilong Zhang,** Rebecca Krouse, Doug Kelkhoff, Matthias Trampisch, Eric Nantz	NA	Ongoing
Requirements/tests workstream set objectives + milestones	**Keaven Anderson**, Nate Mockler, **Tilo Blenk.**	NA	Ongoing
Look out for the white paper for review. Share the white paper and thinking within your companies and any feedback (agreement or disagreement) is useful for the team to know.	Andy Nicholls to distribute in the next 2 weeks for **ALL to review and discuss in your companies.**	Feb 2020	Ongoing
Collect examples of audit processes in pharma to demonstrate the non open source arena to see if this can be used by R open Source.	Doug & Noam **(All)**	NA	Ongoing