

R “Validation” Hub Meeting

02 October 2018

Attendees:

Andy Nicholls (GSK) – Meeting Chair, PSI AIMS SIG
Lyn Taylor (PRA) – Meeting Secretary, PSI AIMS SIG
Bella Feng (Amgen)
Bob Engle (Biogen)
Chris Toffis (Syne qua non) – PSI AIMS SIG
Craig Mcilloney (PPD) – PSI AIMS SIG
Dan Boisvert (Biogen)
Doug Kelkhoff (Roche)
Ed Lauzier (Merck and Co.)
Greg Cicconetti (Abbvie) – ASA BIOP Software
Eric Nantz (Eli Lilly)
Huilei Xu (Novartis)
John Sims (Pfizer)
Jules Hernandez-Sanchez (Roche) – PSI AIMS SIG
Juliane Manitz (Merck Serono)
Keaven Anderson (Merck) – ASA BIOP Software
Kieran Martin (Roche)
Markus Elze (Roche) – PSI AIMS SIG
Martin Gregory (Merck KGaA)
Mat Soukup (FDA CDER)
Min Lee (Amgen) - TransCelerate
Nate Mockler (Biogen)
Paul Schuette (FDA CDER)
Reinhold Koch (Roche)
Rinki Jajoo (Merck)
Satish Murthy (J&J)
Thomas Drgon (FDA)
Xiao Ni (Novartis)
Yilong Zhang (Merck)

Previous Action Items

Action Item	Assigned team member(s)	Deadline	Status
None – 1 st meeting			

R “Validation” Hub Meeting

02 October 2018

Discussion

Andy Nicholls introduced PSI AIMS and everyone on the call introduced themselves.

The following slides were presented, and discussion detailed below.



R Validation Hub
Kickoff Slides.pptx

A consensus from the R/pharma Conference was that using Base R for regulatory work is now accepted. However, additional packages would require more documentation before they could be used for submission work.

Thomas Drgon said that from his perspective there is still a dichotomy between using R for research, and using R to facilitate regulatory decisions. For research, using R is to generate an insight for yourself to further your research. Hence, trying different software and figuring out what works and what doesn't is welcome. However, when using software that would lead to regulatory decisions, the criteria of what can be used may be different. Some R packages exist within the FDA which could be used for regulatory decisions. However, Paul Schuette identified that these packages in his experience have not all used full scale software development lifecycle procedures.

Andy Nicholls, identified that one of the challenges we will face is the many discussions and statements of what we mean by validation. For those of us who want to use R packages downloaded from cran for statistical analysis that may become part of a submission, this is our most concern. The packages used for this, need to have some process so that they are checked to ensure they are doing what they supposed to do. This could be each person doing their own checks but we need some sort of process, not necessarily validation, but we need to have confidence that packages do what we expected to do. “Validation” may not be the right word but instead R “Qualification” or “documentation of use”.

The PSI AIMS SIG idea for an online “Validation” Hub was described which would

- Standardise the packages we use
- Provide links to useful QA information
- Be a platform for sharing tests
- Enable statistical discussion – for example notes regarding comparisons between analysis in R vs SAS
- Lead to a more consistent approach to R validation (and open source in general)
- Be free to use

Ed Lauzier (Merck and Co.) – Suggested that it would be good to have a validated R repo, that could be used and remove the concern about using the wrong version of a package from CRAN. CRAN is wide

R “Validation” Hub Meeting

02 October 2018

open and the versions change frequently, but if we have a CRAN repo or a snapshot of the version of the Code which has been certified, then this would be really powerful moving forward.

Martin Gregory (Merck KGaA) – Confirmed that this is one of the things Merck have been doing to store a local copy of the versions of packages being used which is especially important if you are going to store tests.

Kieran Martin (Roche) – Agreed that he liked the idea of a CRAN snapshot, but you might want to have both. This enables you to have packages that aren't complete enough to exist on the snapshot but may still want to be used (such as older packages which are not kept in line with new testing. Therefore a mix of both package meeting set criteria and ones that don't but still have some valid use would be useful.

Min Lee (Amgen) – Confirmed they have looked at this for a couple of years. If you look at package as a whole, validation is only 1 point in time. Given R packages have a dependency on each other, packages can break and then be hard to rely upon in 1 static environment. Therefore we need to think about how the packages evolve over time. Not just developing a test for packages but develop an environment as a whole that takes the lifecycle of packages and dependencies. We can use tools and processes to take that into account, to look at validation in terms of an environment perspective. Min introduced the TransCelerate program to which she is a member. This is as an Industry consortium with initial approval to go forward with an industry collaborated effort to contribute resources to look at the tooling and processes to actually develop an environment for R package validation.

Juliane Manitz (Merck Serono)– added that for packages with complex mathematics we may only be using certain parts, hence we don't need every package fully tested just the parts we use. So perhaps a weighted network of evidence, where we bring in certain parts that are needed. To do this we'd need to define objective rules. Every R user has a certain set of packages that they trust, likely because they have previous experience using them. Therefore, if we can decide the steps / rules to convince us that a package is trustworthy, then that could be a way forward.

Andy Nicholls – Confirmed that AIMS as a group got to that opinion also, that trying to validate R didn't make sense and it would be a huge challenge, instead guidance on how you can go about qualifying R packages are OK to use appears to be a better approach. For example, listing what to test and what dependencies don't need testing would help. Getting tests for packages is a step but it's about providing a recipe such that people are happy to use R comfortable with the confidence that the software will do what we expect it to do.

Greg Cicconetti (Abbvie) and Keaven Anderson (Merck) are both members of the ASA Biopharmaceutical Section (BIOP) statistical software working group. The group was initiated over the summer 2018 to look at software in the industry with an interest in “validating” R. In addition to testing, Keaven stated they could also do Video training in use of R.

Andy Nicholls – highlighted that it's great to have so many people on the call from different companies and the ASA BIOP and Transcelerate programs however we also need contacts from EMEA / PDMA and

R “Validation” Hub Meeting

02 October 2018

volunteers to contribute from all companies. We need to spread the word of the project and collaborate together to work towards the same aim.

A GitHub page was created by Reinhold Koch (Roche) which we can use to store information about members of the collaboration, definitions, the validation framework (our interpretation of what is required for R packages), and package tests and the risk factors to consider about R packages. In addition we can supply links to other sources for information. This can act as a single page source of information.

Suggest starting with the GitHub page, but as a group we can move in wider directions such as the R environment as a whole and extension to a repository for packages that meet certain standards

All to think about if we are on the right track to meeting the requirements of the industry. Please provide any ideas or concerns about the future plans. We need volunteers to collaborate on different parts of the project.

All to decide which forum they would prefer us to use for future discussion:

Please vote here: <https://www.surveymonkey.co.uk/r/MTL3YRJ>

Actions

Action Item	Assigned team member(s)	Deadline	Status
Set up meeting for 1 months time	Lyn Taylor (PRA)	5 th Oct	Open
To think about how you want to shape the future of the project and what you can contribute so we can discuss at the next meeting	All	25 Oct	Open