**Attendees:**

|  |  |
| --- | --- |
| **Team Member** | **Present at meeting** |
| Craig Mcilloney (PPD) | √ |
| Lyn Taylor (PRA) | √ |
| Chris Toffis (Amgen) | √ |
| David Inman (GSK) | √ |
| Wilmar Igl (AZ) | √ |
| Yann Robert (Servier) | √ |

**Previous Action Items**

| **Action Item** | **Assigned team member(s)** | **Deadline** | **Status** |
| --- | --- | --- | --- |
| Report back to the group areas in their experience that R does it better than SAS. It can be either “Hints & Tips” in R used for exploratory uses (could be a quick gain) or more “regulatory R use” which would need validation.  Such as:  Hints & Tips: R Shiny – interactive graphics much better than SAS but SAS can do this (awkwardly) & spotfire designed to do this & validated. | All | 30 October 2016 | Received from Wilmar & Yann  -Closed |
| Lyn to set up call, to be held before Christmas (hopefully) but after SAS & Mango presentations | Lyn | 12 November 2016 | Closed |
| Lyn to try to arrange a presentation from SAS and research what SAS are currently developing in the next release (9.5?)  Agreed to compare R vs latest SAS | Lyn | 30 October 2016 | Closed (Not possible as don’t release information) |
| Email Mango to invite them to present to us on their current R development & validation progress | Craig | 30 October 2016 | Closed |
| To send the feedback from the SIG meeting to the group | Craig | 30 October 2016 | Open |

**Agenda topics**

| **Topic/Lead** | **Discussion/Decisions** |
| --- | --- |
| All | * What does R do better than SAS?   Wilmar:  -parallel computing  -big data  -genetic analysis  -transparency of statistical models  -innovation, state-of-the-art methods  Yann:  - R more intuitive (especially for new biostatistician who practice R more than SAS at school) and because SAS language is very specific  - Simulations done faster on R (even without parallel computation)  - Parallel computation (we do not try that on SAS)  - Specific methodologies like Bayesian (because not possible with our SAS version) or machine learning (Random Forest…) and all the new statistical methodologies implement in R faster than SAS  - Graph nicer (as compare as SAS) , more flexible and with interactive possibilities using Rshiny  Chris: Spaghetti plots,  Craig: Perhaps we need to write a beginners guide/ basic introduction article in SPIN, including where is R useful? Include why to use R, not necessarily the syntax  Wilmar: Good to produce same analyses done in SAS & R (basic analysis) – Read in, transform variable & do basic regression (Demo).  Lyn: Good to include basic syntax for a SAS user to get data in & do analysis |
| Yann | Our IT team give me more information about Rstudio pro version (the commercial one).   * It allow to use different version of R, to run multiples sessions and to share project (simultaneous code edition) * It manage authentication with LDAP * It allow to manage load balancing and include a monitoring console to follow use and logs * It include commercial support * For an intensive use this commercial version seems to be recommended and ideally the sever one (for the resource scaling and the centralized/simplified administration) |
| Chris / Wilmar | Chris,  Rstudio in house training at Amgen currently in production  Demonstrating Dynamic/interactive plots and the possibility of QCing an output generated by SAS within R  Wilmar, Useful doc: Differences between Rstudio commercial vs. open-source:  <https://www.rstudio.com/products/rstudio/#Server>  Perhaps create separate article for SPIN on RSTUDIO, server, commericial vs open-source & what it can do more than the basic R. |
| Lyn | Lyn’s spin competition idea: sponsorship for a £50 amazon voucher prize? Ask for people who use R in industry to write about their experience?  May not get much feedback but perhaps may get some junior statisticians interested? |
| Lyn/Craig | Mango solutions scheduled for 21st Nov 11am.   * Get their progress on validation. * They do the work for GSK (David to find out contact & what they use if possible) |
| Lyn | * R-Foundation board (Marc Schwartz) email: Key points below, what are the next steps?   + No software is recommended by the FDA or pre-validated. FDA do use R internally & R used commonly in pre-clinical work.   + SAS’s dominance is historical due to the fact a large body of the end user written analytic code has been tested & validated by end users already.   + End user must engage in internal activities that satisfy the regulatory framework for any software.   + Industry transition to XML based format (CDISC) means less reliance on XPORT formats (& other software can create XPORT anyway)   + Can use R collaboratively with SAS or shift analyses to R?   + List of “Base R plus Recommended Packages” are standard R distribution from CRAN (Copyrighted?) – All other packages are not official releases. But there may be situations where you need specific methods only available in 1 package from another author.   + To use a package:     - 1) determine whether there is sufficient history of use of the package?     - 2) does the author/maintainer of the package have a firm reputation?     - 3) Are they responsive to communications?     - 4) do they seem to use standard software development, patch and version control procedures?     - 5) How good is their documentation and has the methodology used in the package been published in a peer reviewed journal and/or cited widely in journal publications   Group discussion:  Individual reviewers at FDA may use R so have to check for submission if they accept R, or SAS only. May depend on the reviewer.   * ACTION: discuss with MHRA/EMEA (FDA?) their current R usage & how they see things going in future, what do they look for from us to satisfy them that the package used (R package used for the submission) is validated. Craig to ask Anna. * If the package used is not familiar it may hinder the submission as seen as weak – makes the submission harder than just using SAS. What can we do to satisfy them? |
| All / Wilmar (Domino data lab | * Other software:   + Domino Data Lab <https://www.dominodatalab.com> A reproducible cloud computing platform integrating various tools (R, SAS, python) -Domino Data Lab, in addition to free R:   https://www.dominodatalab.com/benefits   * + Sweave?   + Knitr? |
| All | AOB & Next Steps |

**Action Items**

| **Action Item** | **Assigned team member(s)** | **Deadline** | **Status** |
| --- | --- | --- | --- |
| Write SPIN R introduction for industry statisticians (not a technical stance but more Why is learning R a worthwhile effort)  Chris to start by providing some text in this area & share with group | Chris | 24 Dec 2016 | Open |
| Rstudio Introduction for industry statisticians. Wilmar to make a start in this area and share with group | Wilmar | Approx: January 2017 | Open |
| MHRA/EMEA their current usage & how they see things going in future, what do they look for from us to satisfy them that the package used (R package used for the submission) is validated. Craig to ask Anna. | Craig | December 2016 | Open |
| To send the feedback from the SIG meeting to the group | Craig | December 2016 | Open |
| If possible, find out how Mango solutions have been used at GSK | David | December 2016 | Open |