R - Eliminating Fear, Uncertainty and Doubt

Mike K Smith, Pfizer

Global Pharmacometrics

Disclaimer

This talk represents the opinions of the speaker and are not necessarily shared by others at Pfizer Ltd., and specifically, do not necessarily represent Pfizer Quality, IT, Statistics, or Legal department opinions.

Acknowledgements

Ian Francis

Frank Harrell Jr.

Max Kuhn

Rich Pugh

Anthony Rossini

Marc Schwartz

David Smith

Mat Soukup

Fear, Uncertainty, Doubt

Fear, Uncertainty, Doubt

… a sales or marketing strategy of disseminating negative (and vague) information on a competitor's product. (Wikipedia)

VS

WHICH IS BETTER?

…LET'S NOT GO THERE

Let’s play
“True” or “False"

True or False?

Regulatory agencies will only accept statistical analysis and reports prepared using validated systems such as SAS and S-Plus.

TRUE…

Regulatory agencies will only accept statistical analysis and reports prepared using validated systems such as SAS and S-Plus.

…FALSE

"The Food and Drug Administration does not endorse or require use of any specific software…"

- Mat Soukup, FDA (UseR Conference 2007)

FDA Expectations

2002 General Principles of Software Validation - Final

2003 Guidance Part 11: Electronic Records - Final

2007 Guidance on Computerized Systems Used in Clinical Investigations - Final

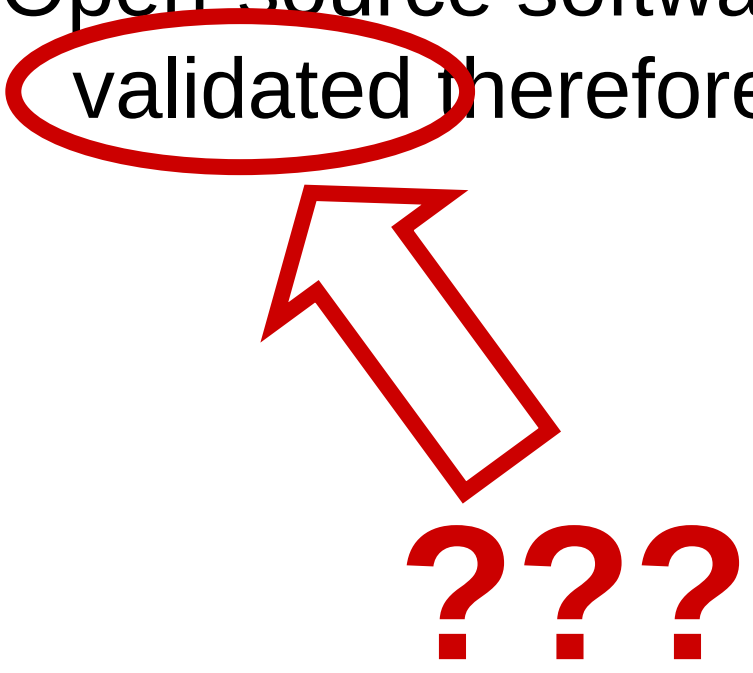
Summary of key points applied to clinical trials presented by:
Bell et al., JSM, 2006

True or False?

Open-source software such as R cannot be validated therefore is not acceptable.

True or False?

Open-source software such as R cannot be validated therefore is not acceptable.



???

What do we mean by “Validation”?

“Establishing documented evidence which provides a high degree of assurance that a specific process will consistently produce a product meeting its predetermined specifications and quality attributes.” – FDA definition.

What do we mean by “Validation”?

“Establishing ***documented evidence*** which provides a high degree of assurance that a specific process will consistently produce a product meeting its predetermined specifications and quality attributes.” – FDA definition.

What do we mean by “Validation”?

“Establishing documented evidence which provides a high degree of assurance that a specific process will consistently produce a product meeting its predetermined specifications and quality attributes.” – FDA definition.

What do we mean by “Validation”?

“Establishing documented evidence which provides a high degree of assurance that a specific process will ***consistently*** produce a product meeting its predetermined specifications and quality attributes.” – FDA definition.

What do we mean by “Validation”?

“Establishing documented evidence which provides a high degree of assurance that a specific process will consistently produce a product meeting its ***predetermined*** specifications and quality attributes.” – FDA definition.

Validation

≠

21 CFR part 11 compliance

21 CFR part 11

21 CFR Part 11 was issued by the FDA to provide regulatory requirements for processes and controls that must be applied to electronic records and electronic signatures.

i.e. output from validated systems.

21 CFR part 11

The guidance can be summarized as follows:

- If there are **predicate rules that require records to be maintained, and these records are managed electronically**, then Part 11 controls apply to those records.
- If there are **predicate rules that require a signature to be applied, and this signature is applied electronically or digitally**, then Part 11 controls apply.

21 CFR part 11

"Records submitted to FDA, under predicate rules in electronic format [are Part 11 records]. However, a record that is ***not itself submitted***, but is used in ***generating a submission***, is not a part 11 record unless it is otherwise required to be maintained under a predicate rule and it is maintained in electronic format."

Good Practice

Installation qualification (IQ)

Is the software installed and maintained according to Software Development Life Cycle (SDLC) processes / SOPs?

Operation qualification (OQ)

Is the software ***consistent*** / stable?

Performance qualification (PQ)

Is the software producing results that are ***credible***, ***reproducible*** and meeting ***agreed*** standards?

And all of this ***documented…***

Good practice

leads to

Validation

What was the question?

Open-source software such as R cannot be validated therefore is not acceptable.

SO… FALSE

YOUR SOPs and practice governing IQ/OQ/PQ will help dictate whether ***YOUR*** use of R falls within GxP.

21 CFR Part 11 doesn't necessarily apply to software systems *per se* but their output.

(See R-FDA document for more details).

True or False?

Since R has been developed by academics and non-commercial organisations who have no stake in the correctness or otherwise of their code, we cannot rely on results being correct.

TRUE… BUT…

Commercial products have their faults too.

Remember Microsoft Excel's Standard Errors?

Is SAS entirely bug-free? Viz: SAS Hot Fixes.

Example: nlme

nlme function is in S-Plus for analysis using nonlinear mixed effects models, written by Doug Bates, Jose Pinheiro…

…The same function appears in R.

CRAN = Comprehensive R Archive Network

nlme: Linear and Nonlinear Mixed Effects Models

Fit and compare Gaussian linear and nonlinear mixed-effects models.

Version:	3.1-90
Priority:	recommended
Depends:	graphics, stats, R (≥ 2.4.0)
Imports:	lattice
Published:	2008-12-30
Author:	Jose Pinheiro, Douglas Bates, Saikat DebRoy, Deepayan Sarkar, the R Core team.
Maintainer:	R-core <R-core at R-project.org>
License:	GPL (≥ 2)
Citation:	nlme citation info
In views:	ChemPhys, Econometrics, Environmetrics, SocialSciences
CRAN checks:	nlme results

Downloads:

Package source:	nlme_3.1-90.tar.gz
MacOS X binary:	nlme_3.1-90.tgz
Windows binary:	nlme_3.1-90.zip
Reference manual:	nlme.pdf
News/ChangeLog:	ChangeLog
Old sources:	nlme archive

Reverse dependencies:

Reverse depends:	AdaptFit, FunNet, GRRGI, HydroMe, JM, LoopAnalyst, PKtools, RHmm, SemiPar, SimHap, agce, arrayImpute, assist, bear, bs, calib, cem, drc, eba, far, gbs, ig, kinship, lmeSplines, longRPart, meboot, mixlow, mmlcr, multilevel, nlmeODE, nlrwr, pheno, picante, plm, portfolio, psychometric, ramps, randomLCA
Reverse imports:	ape, mgcv, psychometric, urca
Reverse suggests:	AER, Design, Epi, MBESS, R2HTML, RLRsim, Rcmdr, RcmdrPlugin.IPSUR, VR, aod, bbmle, contrast, dlmap, fda, gmodels, mgcv, multcomp, npmlreg, pgirmess, primer, psyphy

True or False?

R functions could easily be edited by any user to change their algorithms and so cannot be validated.

TRUE…

Users ***CAN*** alter functions to “mask” those from standard packages.

- You DO get a warning however.

…AND FALSE

Validation argument though comes back to GxP and reproducibility.

- If ***YOUR*** functions follow SDLC and are reproducible then their use can be “validated”.

True or False?

Any work performed in R must be validated by replication using other "approved" software before submission.

True or False?

Any work performed in R must be validated by replication using other "approved" software before submission.

FALSE

IF you have followed GxP, SDLC, IQ/OQ/PQ processes, SOPs etc. for R installation and use…

THEN there should be no need to reproduce work.

REMEMBER

"The Food and Drug Administration does not endorse or require use of any specific software…"

- Mat Soukup, FDA (UseR Conference 2007)

ASIDE…

SAS have released experimental procedures for Bayesian analysis and PROC MCMC.

Do we use PROC MCMC or WinBUGS / OpenBUGS as the benchmark?

True or False?

R has never been used for business critical work at the FDA.

FALSE

Statistical review (meta-analysis) of Avandia (rosiglitazone) at FDA Advisory Committee by Joy Mele.

R at the FDA

Mat Soukup reports in his presentation that there is a growing number of statisticians at FDA who use R.

- R software is “approved” for use by FDA IT organisation.

R used by other disciplines

Xpose suite of diagnostic tools for Population Pharmacokinetics / NONMEM models.

- Written in R
- “Standard” tool for NONMEM diagnostics
- Used by Pharmacometricians / Clinical Pharmacologists.
- Quite likely some of them at regulatory agencies. (e.g. Sweden)

True or False?

R can only be used for non-regulatory work and so is best used in the preclinical and nonclinical space.

FALSE

GxP / SOP / SDLC overhead ***MAY*** mean that it is ***EASIER*** to use it for non-regulatory, preclinical and non-clinical work.

- R does have some very good packages for use in the preclinical / non-clinical arena.

BUT R ***CAN*** be used for regulatory work.

True or False?

We have nobody in-house who understands R well enough to QC code. Therefore it should not be used.

TRUE

If you have nobody in-house to QC your code, then you probably shouldn't be using R for regulatory submission.

BUT the same goes for any other package including commercial packages e.g. S-Plus, GENSTAT, STATA, MINITAB etc.

…Same issue with SAS

SAS Macro code, SQL, IML, ODS…

There's good money to be made in the SAS programming contractor / consultancy world…

- Programming for job security

Caveats

R can easily be downloaded an loaded onto computers within an organisation.

- It's FREE as in beer, and as in speech.

- Packages can be installed directly from within R.

This can lead to issues about version control.

Conclusions

IF you have a good SDLC process, SOPs, etc. controlling roll out of software, coding standards, QC of code, scripting etc.

THEN there are very few reasons why analysis performed in R wouldn’t be accepted by regulatory agencies.

R and the Pharma Industry

Benefactors / Supporting Institutions

- Merck
- Astra Zeneca
- Baxter
- Boehringer Ingelheim

Used at

- Pfizer
- Novartis
- Roche
- GSK
- AstraZeneca
- Pharmerit
- NovoNordisk
- PRI
- Sanofi Aventis

References

Using R: Perspectives of a FDA statistical reviewer (Mat Soukup)

Open Source Statistical Software (OS3) in Pharma Development: a case study with R (Anthony Rossini)

R for clinical trial reporting: reproducible research, quality and validation (Frank Harrell)

Times “R” a changing: FDA perspectives on use of “Open Source” (Sue Bell)

FDA meta-analysis of Avandia (rosiglitazone) NDA 21-071 Supplement 022 (Joy Mele)

R: Regulatory compliance and validation issues. A guidance document for the use of R in regulated clinical trial environments (The R Foundation for Statistical Computing)

List of R Foundation members and supporters

Analyzing clinical trial data for FDA submissions with R (David Smith)