Some Challenges of Using *R* in a Regulatory Environment

Jae Brodsky

Office of Biostatistics, OTS, CDER, FDA

The Big Question

How well can *R* be used as a statistical tool in a large federal regulatory agency such as the FDA?

Current *R* Usage at the FDA

- FDA offices use *R* on a daily basis
  - includes Office of Biostatistics
- FDA scientists have written *R* packages for other scientists' use (FDA or non-FDA)
- FDA does not endorse or require the use of any specific software for statistical analysis

Regulations and Guidelines for Regulatory Use

- Code of Federal Regulations Title 21: FDA
  - Part 11: Electronic records and electronic signatures
- General Principles of Software Validation: Final Guidance for Industry and FDA Staff (2002)
- Guidance Part 11: Electronic Records (Final) (2003)
- Guidance for Industry: Computerized Systems Used in Clinical Investigations (2007)
- ICH E6: Good Clinical Practice Consolidated Guideline
- ICH E9: Statistical Principles for Clinical Trials

Code of Federal Regulations

21 CFR 11 deals with electronic records and signatures that are "trustworthy, reliable and generally equivalent to paper records."

- Qualification, verification and validation
- *R* is only part of a larger system that is verified
- *R* Foundation compliance document (2008): http://www.r-project.org/doc/R-FDA.pdf

This poster reflects the views of the author and should not be construed to represent FDA's views or policies.

Regulations and Guidances: What Applies to *R* at the FDA

Regulations state the specifics of what is legally required and enforceable. Guidances describe the agency's current position and are recommendations. 21 CFR Part 11 regulates electronic records and their storage, not the software that is used to generate reports. There are no regulations that restrict the use of open source software (including *R*) at the FDA. However, there are concerns for the FDA beyond CFR Title 21.

Current FDA *R* Setup

Only base and recommended packages are fully supported by the R core. Code or functions outside of these packages are used at the reviewer's own risk.

- Base: *datasets, graphics, grDevices, grid, methods, splines, stats, stats4, tcltk, tools, utils*
- All results from data should be reproducible independent of software used.
- FDA is not responsible if results have discrepancies.

Sponsors may use R in their submissions. Data must be submitted in *xpt* format. Reviewers would like to know:

- which *R* functions were used
- how they were accessed
- that the results are accurate, reliable, and consistent

Regulatory Agency Needs

Freeware is good in theory for a government agency, but comes with several issues:

- No tech/customer support
- How should the FDA provide support to *R* users?

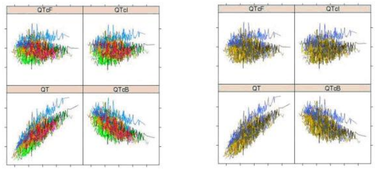
The FDA has used *SAS* for 30+ years. FDA groups are using *SAS* and *R* concurrently. Legacy code is generally dealt with by calling *SAS* from *R*.

- Is this an effective or efficient way to handle old *SAS* code?
- *R* users must keep up with *SAS* code changes, and vice versa.

The FDA does not endorse any statistical software. Can we be innovative with *R* without appearing as though we "endorse" *R*?

Further Regulatory Issues That Impact *R* Usage at the FDA

As a government agency, the FDA must comply with additional regulations that do not impact *R*. Section 508 of the Rehabilitation Act of 1973 (29 U.S.C. 794d) established guidelines for ensuring that the federal government's electronic information is accessible to people with disabilities, including blindness, deafness, colorblindness, epilepsy, and reduced motor function. Section 508 includes guidelines for luminosity or brightness, use of color, font size, text alternatives for figures or plots, and text-only alternatives. All reports released online by the FDA should be 508 compliant, including reports from the Office of Biostatistics.



Example from an FDA report: Left: a plot made with the default rainbow color palette. This figure is not 508 compliant because some information is lost to colorblind people. Right: the red-green colorblind simulation. It is not possible to determine which lines correspond to individual subjects when the color palette is reduced.

Would the FDA need to modify or write its own version of popular *R* packages to meet additional guidelines? Section 508 compliance is not mandated for *R*, and the FDA cannot ask the *R* community to automatically make packages 508 compliant.

My Personal Experience as an *R* User at the FDA

The QT interval is a biomarker for cardiac repolarization. The QT Interdisciplinary Review Team (IRT) reviews Thorough QT (TQT) studies to investigate QT interval lengthening. The QT IRT consists of statisticians, clinical pharmacologists, and clinicians.

I am the only *R* user in the stats group (4 statisticians). I use:

- *SAS* for reviews (for consistency with other QT IRT stats reviewers)
- *R* for research
- *R* for any graphics outside reviews

The QT IRT clinical pharmacologists use *R* for their reviews. The *R* package *QT* was written specifically for the QT IRT.

- *QT* calls the stats group's *SAS* macros for areas in the clin pharm portion of the review
- other areas of the clin pharm review are done in *R*
- original author of *QT* no longer at the FDA

My future plans: to write an *R* package for QT IRT use that covers all parts of a TQT review.

Timeline: ?

In my experience, *R* use at the FDA is completely acceptable and has not caused any problems.

References

[1] Title 21 CFR Part 11 Section 11.1 (a)

[2] General Principles of Software Validation; Final Guidance for Industry and FDA Staff

[3] Guidance Part 11: Electronic Records (Final)

[4] Guidance for Industry: Computerized Systems Used in Clinical Investigations

[5] ICH E6(R1): Good Clinical Practices

[6] ICH E9: Statistical Principles for Clinical Trials

[7] *R*: Regulatory Compliance and Validation Issues: A Guidance Document for the Use of *R* in Regulated Clinical Trial Environments:

[8] Section 508 of the Rehabilitation Act of 1973 (29 U.S.C. 794d):