Using R: Perspectives of a FDA Statistical ReviewER

Mat Soukup, Ph.D.
Mat.Soukup@fda.hhs.gov

FDA/CDER/OTS/OB/DB3

Disclaimer

- The views expressed in this presentation are those of the presenter and must not be taken to represent policy or guidance on behalf of the Food and Drug Administration.
- The Food and Drug Administration does not endorse or require use of any specific software including those software mentioned in the presentation.

Outline

Guidance: *Computerized Systems Used in Clinical Trials*

In March 1997, FDA issued 21 CFR part 11, which provides criteria for acceptance by FDA, under certain circumstances, of electronic records, electronic signatures, and handwritten signatures executed to electronic records as equivalent to paper records and handwritten signatures executed on paper. After the effective date of 21 CFR part 11, significant concerns regarding the interpretation and implementation of part 11 were raised by both FDA and industry. As a result, we decided to reexamine 21 CFR part 11 with the possibility of proposing additional rulemaking, and exercising enforcement discretion regarding enforcement of certain part 11 requirements in the interim.

Final Draft May 2007

Regulations vs. Guidances

"FDA's guidance documents do not establish legal enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, not required."

What do the Regulations and Guidances Say?

- No regulations prohibit the use of R for statistical analysis.
- Guidances DO provide 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
  - Petullo *et al.*, DIA, 2007

Outline

A Regulatory Submission Using R?

The Agency's primary objective is to determine if the drug/biologic/device is safe and effective.

- Organizations SOP should prevent user-related errors.
- Use of R functions without proper validation is at the organizations risk.
- Results should be reproducible and independent of the software used to derive them.
- When results are not reproducible; FDA is not responsible for rectifying discrepant results.

If/When a Submission Relies on R

Reviewer Expectations/Requests

1. What R functions are used and where do they reside (base vs. user-contributed packages)?
2. Has the R function been properly validated?
   - Can the validation tests be reproduced?
   - Is there any certification of the validation test?
   - Are there any known data structures which can potentially alter results?

The reviewer wants to be assured that the results are accurate, reliable, and consistent.

Guidance: *General Principles of Software Validation*

6.3. VALIDATION OF OFF-THE-SHELF SOFTWARE AND AUTOMATED EQUIPMENT

Most of the automated equipment and systems used by device manufacturers are supplied by third-party vendors and are purchased off-the-shelf (OTS). The device manufacturer is responsible for ensuring that the product development methodologies used by the OTS software developer are appropriate and sufficient for the device manufacturer's intended use of that OTS software. For OTS software and equipment, the device manufacturer may or may not have access to the vendor's software validation documentation. If the vendor can provide information about their system requirements, software requirements, validation process, and the results of their validation, the medical device manufacturer can use that information as a beginning point for their required validation documentation. The vendor's life cycle documentation, such as testing protocols and results, source code, design specification, and requirements specification, can be useful in establishing that the software has been validated. However, such documentation is frequently not available from commercial equipment vendors, or the vendor may refuse to share their proprietary information.

Materials to Aid FDA Reviewers

Contributions from the R Community

- Website which contains details of the validation.
- Repository of all versions of R which include the following.
  - All executable or installation files of the R program.
  - Listing of all packages which have been properly validated in terms of providing accurate, reliable, and consistent results.

With access to the above information, this *exceeds* current regulation standards!

Outline

Installing R on FDA-issued Computers

- Non-standard software are not allowed to be installed on government computers without IT approval.
- 'Business case' is needed to gain IT approval; can be a slow process.
- R was officially approved as a non-standard software in 2004; approval did not include any future R releases.
- In early 2007, R was officially granted a constrained "version-less" approval; requires small group of adopters at early release.

Uses of R within FDA

- Who uses it?
  - Statisticians; mostly recent graduates, some conveRts
  - Potentially others with quantitative backgrounds (e.g. pharmacometricians in CDER)
  - Internal course will be offered which has potential to increase numbers of UseRs.
  - My experience: The S language sells itself; it just requires a teacheR!
- What is R used for?
  - Product review (NDA/BLA/PMA) to assess safety and efficacy.
  - Research/Simulations
  - Advisory Committee: e.g. statistical review of rosiglitazone by Joy Mele, M.S.

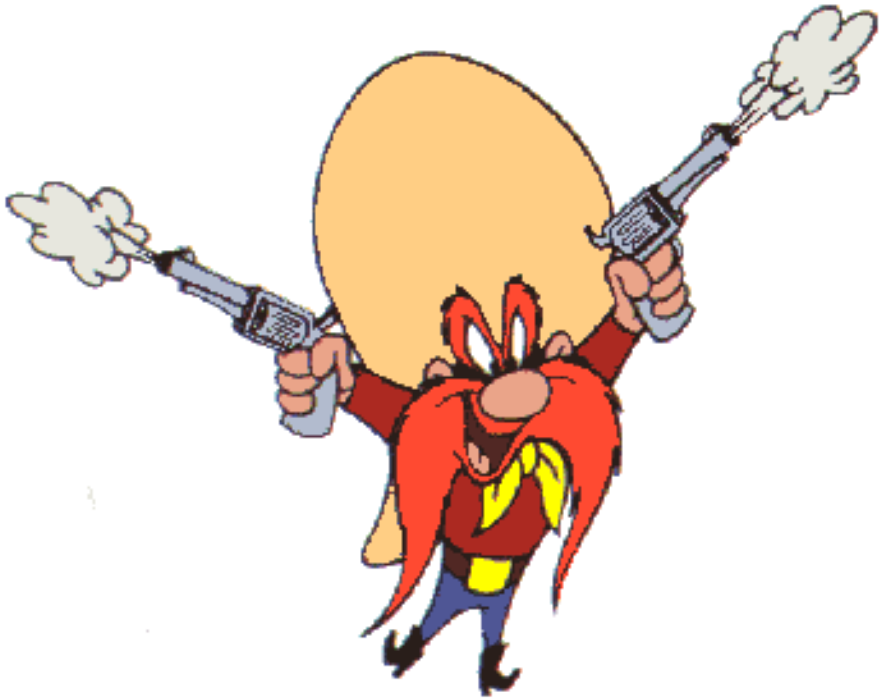
My RevieweR Personality

The RegulatoR

My RevieweR Personality

The RegulatoR

The UseR

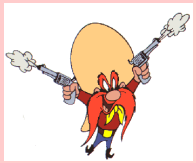


My RegulatoR Personality

- Perfectionist - **need/require** accuracy of data derived conclusions.
- Conservative - tradition and past precedence have strong role on actions.
- Slow to Change - not resistant, just late adopters of new trends.

Dominant when findings of safety and efficacy are critical (e.g. Advisory Committees). **R *may* be my software of choice.**

My UseR Personality



- Perfectionist - **desire** for accuracy of data derived conclusions.
- Liberal - open to new approaches which can be used as sensitivity analysis.
- Adventurous - see how the data tells the story; not everything can be pre-specified; use graphical depictions!

Dominant when important results are presented and reproducible (e.g. NDA efficacy). **R *is* my software of choice.**

Outline

Potential Implementations of R in the Future

- With a shift to understanding safety data, there is a greater need for active assessment of the data rather than pre-specifed reporting of data results; paradigm of active data analysis. R excels at such a paradigm!
- Sweave like statistical/clinical standard review generation; improves efficiency for generating standard set of tables and figures with more time available for active data analysis.
- Open Toolbox - open source environment in which analysis tools can be developed which can inter-play with multiple software vendors.

Closing Remarks

- Regulatory Conclusions
  - Regulations do NOT prevent use of R.
  - However, validation, documentation, and accountability are required of sponsors relying on R.
  - A public listing of validated R functions would
    - be a great benefit to FDA reviewers and sponsoring companies.
    - *exceed* current standards - R sets the standard!
- FDA RevieweR/UseR Perspectives
  - The Agency's primary objective is to determine safety and efficacy which should be independent of the software used to derive the results.
  - For data driven conclusions there is a need for accuracy, reliability, and consistency.
  - Sponsor's conclusions need to be reproducible - should be independent of software.

URL's to Guidances

- Guidance: "Part 11, Electronic Records; Electronic Signatures Scope and Application" issued August 2003 available at `http://www.fda.gov/Cder/guidance/5667fnl.pdf`.
- "General Principles of Software Validation: Final Guidance for Industry and FDA Staff" issued January 2002 available at `http://www.fda.gov/cdrh/comp/guidance/938.html`.
- "Computerized Systems Used in Clinical Investigations" - Final Guidance dated May 2007 available at `http://www.fda.gov/cder/guidance/7359fnl.pdf`.
- "Off-The-Shelf Software Use in Medical Device" dated September 1999 available at `http://www.fda.gov/cdrh/ode/guidance/585.html`.