

[Unfulfilled?] Potential of R in Clinical Research

EDUARD PARSADANYAN

Biostatistics & Statistical Programming

29 OCT 2015, BERLIN

AGENDA

- **What is Clinical Research?** (brief intro and terminology)
- **What is the potential of R?** (in variously regulated environments)
- **What needs to be considered for using R in regulated environments?**
 - Security breaches (case-study with a Shiny App for bioequivalence)
 - Industry standards (for computerized systems and data deliverables)
 - Audits (how to prove that words meet actions)

WHAT IS CLINICAL RESEARCH?



Clinical research is a branch of healthcare science that **determines the safety and effectiveness of** drug/device/biologic **products** and treatment regimens **intended for human use.**

These may be **used for prevention, treatment, diagnosis or for relieving symptoms of a disease.**

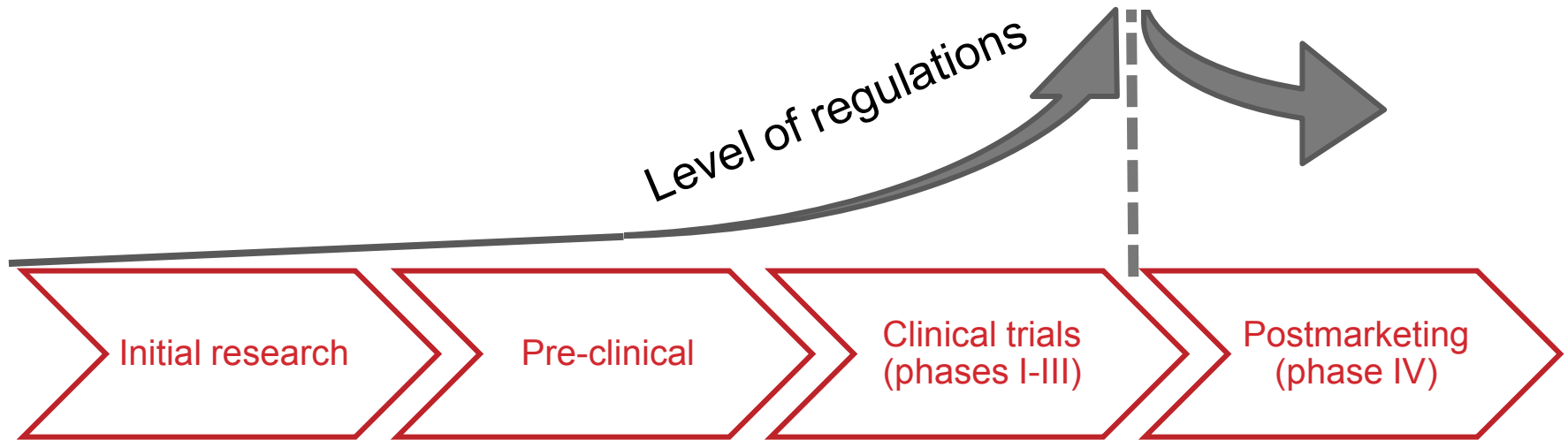
WHAT IS CLINICAL RESEARCH?



Main actor are

- **Sponsors** (e.g. pharma company that own the molecule)
- **Subcontractors** (e.g. contract research organizations – CROs)
- **Regulators** (government agencies like FDA and EMA)

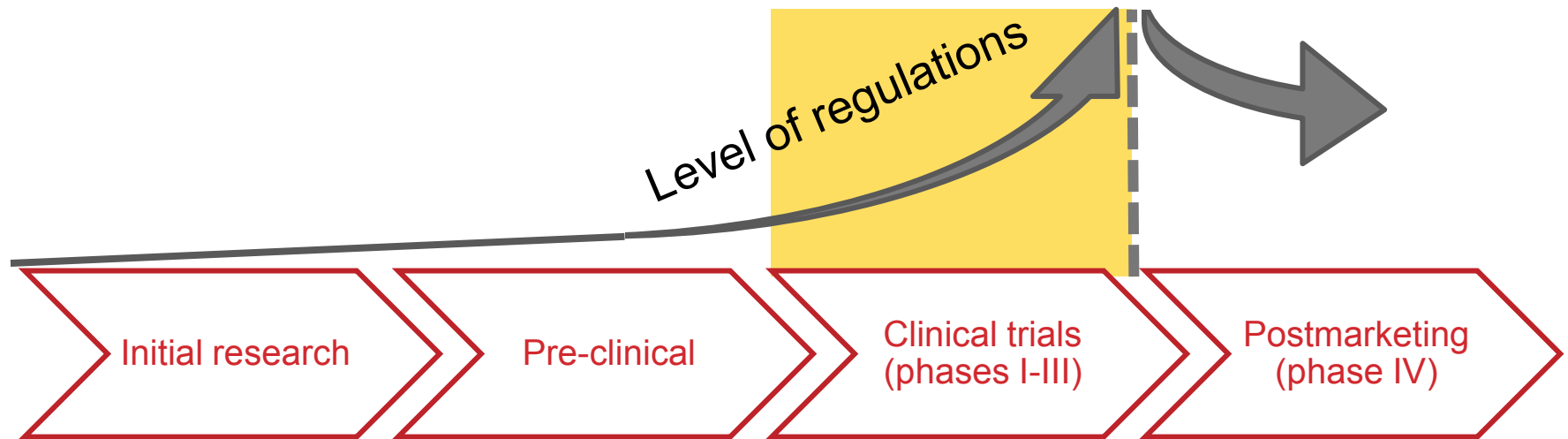
WHAT IS THE POTENTIAL OF R?



NB! Lack of regulations does not mean absence of any standards

Key points: Usage of R is an invers function of regulatory scrutiny
Limitations are confused with prohibition

WHAT IS THE POTENTIAL OF R?



NB! Lack of regulations does not mean absence of any standards

Key points: Usage of R is an invers function of regulatory scrutiny
Limitations are confused with prohibition

CONSIDERATIONS FOR USING R IN REGULATED ENVIRONMENT

- **Security breaches**

In most cases security issues are not directly related to R but sometimes they do.

- **Industry standards**

- For computerized systems (possibility to use R in principle)
- For deliverables (electronic submissions to FDA)

- **Audits**

Sponsors, government agencies, independent parties.

CASE STUDY

Shiny-app for bioequivalence studies by Andrey Ogurtsov:

statist.shinyapps.io/bioeq_en

FDA defines bioequivalence as:

“the absence of a significant difference in the rate and extent to which the active ingredient ... in pharmaceutical equivalents [i.e. generics] becomes available at the site of drug action when administered at the same molar dose under similar conditions in an appropriately designed study”.

CASE STUDY

Shiny-app by Andrey Ogurtsov: statist.shinyapps.io/bioeq_en

The screenshot shows a web browser window displaying the 'Bioequivalence v0.3' Shiny application. The browser's address bar shows the URL 'https://statist.shinyapps.io/bioeq_en'. The application header includes the 'shinyapps.io' logo and 'Powered by R Studio'. A navigation bar contains tabs for 'Bioequivalence v0.3', 'Data upload' (which is highlighted), 'Source data', 'Plots', and 'Results'. The 'Data upload' section is active, showing a file upload interface on the left with a 'Choose .txt file' button, a file named 'testdata.txt' with a Russian label 'Обзор...', and an 'Upload complete' button. Below this, it specifies 'Tab separated data with comma as decimal point'. The main content area is titled 'Data upload' and explains that the app computes results for a standard 2x2 crossover design. It lists the required dataset columns: 'subj' (randomization number), 'seq' (sequence 1-RT, 2-TR), 'prd' (period), 'drug' (test or reference formulation), 'time' (time of blood sampling), and 'conc' (concentration of analyte). It also provides instructions on how to download a demo dataset and how to handle undetectable concentrations (0 for first time points, NA for last points).

shinyapps.io Powered by R Studio

Bioequivalence v0.3 Data upload Source data Plots Results ...

Choose .txt file

Обзор... testdata.txt

Upload complete

Tab separated data with comma as decimal point

Data upload

This app computes main results required for bioequivalence evaluation (standart 2x2x2 crossover design). Your dataset should contains:

- subj - randomization number;
- seq - sequence (1 - RT, 2 - TR);
- prd - period;
- drug - test (T) or reference (R) formulation;
- time - time of blood sampling;
- conc - concentration of analyte.

Demo dataset can be dowloaded [here](#) (right click - save as...).

Undetectable concentrations for first time points should be entered as 0; for last points - as NA (missed values).

CASE STUDY

Shiny-app by Andrey Ogurtsov: statist.shinyapps.io/bioeq_en

shinyapps.io Powered by R Studio

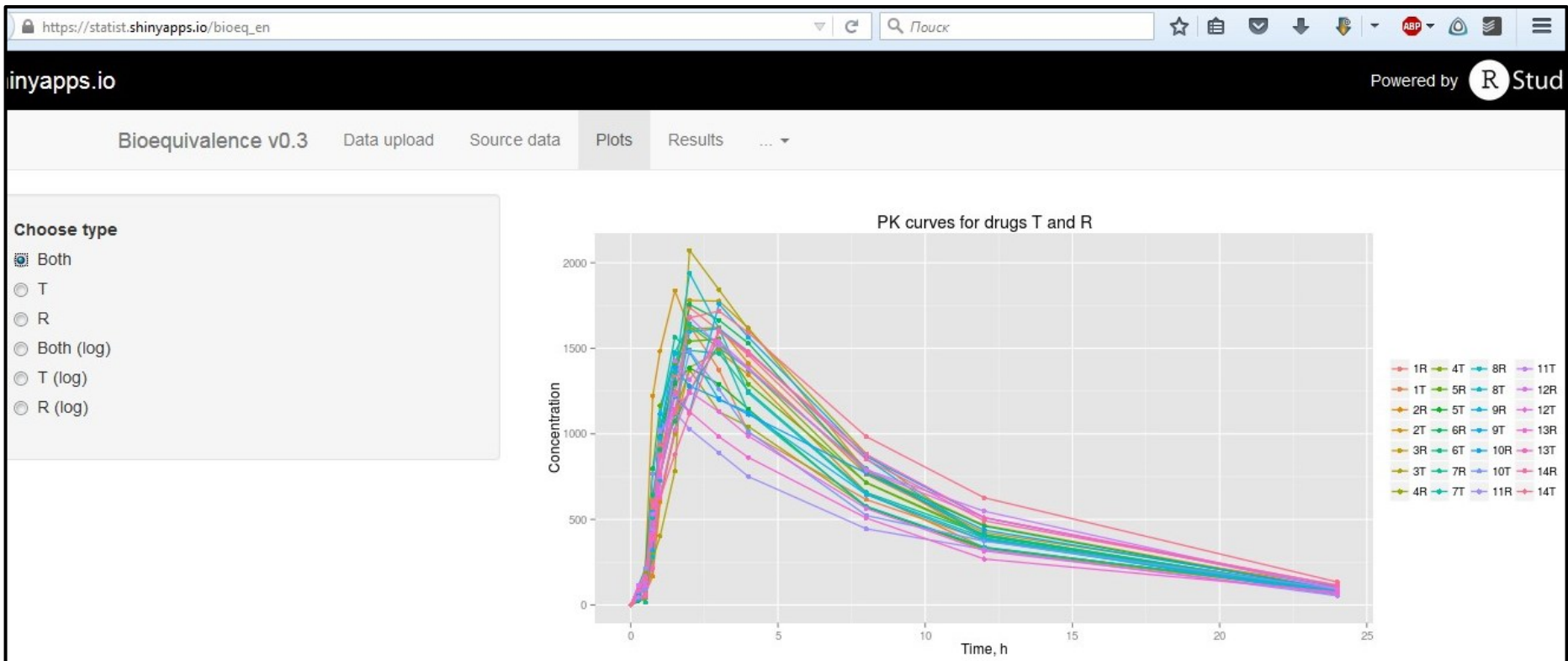
Bioequivalence v0.3 Data upload Source data Plots Results ...

Show 25 entries Search:

subj	seq	prd	drug	time	conc	subj2
1	2	2	R	0.00	0.0	1R
1	2	2	R	0.25	36.1	1R
1	2	2	R	0.50	125.0	1R
1	2	2	R	0.75	567.0	1R
1	2	2	R	1.00	932.0	1R
1	2	2	R	1.50	1343.0	1R
1	2	2	R	2.00	1739.0	1R
1	2	2	R	3.00	1604.0	1R
1	2	2	R	4.00	1460.0	1R
1	2	2	R	8.00	797.0	1R

CASE STUDY

Shiny-app by Andrey Ogurtsov: statist.shinyapps.io/bioeq_en



CASE STUDY

Shiny-app by Andrey Ogurtsov: statist.shinyapps.io/bioeq_en

shinyapps.io

Powered by R Studio

Bioequivalence v0.3

Data upload

Source data

Plots

Results

...

Cmax

	Value	Std.Error	DF	t-value	p-value
(Intercept)	7.36	0.05	12.00	142.87	0.00
drugT	0.06	0.05	12.00	1.27	0.23
prd2	-0.06	0.05	12.00	-1.31	0.21
seq2	0.00	0.06	12.00	0.03	0.98

AUC(0-t)

	Value	Std.Error	DF	t-value	p-value
(Intercept)	9.18	0.12	12.00	74.61	0.00
drugT	-0.02	0.12	12.00	-0.15	0.88
prd2	0.04	0.12	12.00	0.30	0.77
seq2	0.13	0.12	12.00	1.07	0.31

90% CI

	Lower limit of 90% CI, %	T/R ratio, %	Upper limit of 90% CI, %
Cmax	97.59	106.23	115.64
AUC(0-t)	78.81	98.13	122.18

CONSIDERATIONS FOR USING R IN REGULATED ENVIRONMENT

- **Security breaches**

In most cases security issues are not directly related to R but sometimes they do.

- **Industry standards**

- For computerized systems (possibility to use R in principle)
- For deliverables (electronic submissions to FDA)

- **Audits**

Sponsors, government agencies, independent parties.

INDUSTRY STANDARDS

- **Documents collectively referred to as GxP:**
 - 21 CFR Part 11 - Electronic Records; Electronic Signatures
 - Guidance for Industry: Part 11, Electronic Records; Electronic Signatures - Scope and Application
 - 21 CFR Part 58 - Good Laboratory Practice (GLP)
 - 21 CFR Part 312 - Good Clinical Practice (GCP)
 - 21 CFR Part 210 - Current Good Manufacturing Practice (cGMP)
 - ICH E6 - Good Clinical Practice Consolidated Guideline
- **Principal software guidance documents:**
 - Guidance for Industry - Computerized Systems Used in Clinical Investigations (2007)
 - General Principles of Software Validation; Final Guidance for Industry and FDA Staff (2002)
- **Principal statistical guideline documents:**
 - ICH E9 - Statistical Principles for Clinical Trials
 - Guidance for Industry and FDA Staff - Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials (2010)

INDUSTRY STANDARDS

Key points

- Standards do not specify which statistical software can or should be used
- There is a distinct difference between data collection and storage (21 CFR Part 11 - Electronic Records) vs data processing & reporting (General Principles of Software Validation)
- Most standards for computerized systems are already met in established companies
- R installation need to be validated

The R Foundation prepared a document “R: Regulatory Compliance and Validation Issues” in Dec 2014: www.r-project.org/doc/R-FDA.pdf

Revolutionanalytics blog article about R in FDA:
blog.revolutionanalytics.com/2012/06/fda-r-ok.html

INDUSTRY STANDARDS contd.

- **FDA expects data to be submitted in electronic ways**
- **There are standards developed by CDISC** – Clinical Data Interchange Standards Consortium: CDASH, SDTM, ADaM etc., some of them are adopted by FDA.
- **FDA requires datasets to be sent in .xpt (Transport File) format** originally developed by SAS Institute. This is an open standard now which is already supported in R.
- **CDISC works towards even more platform independent standards based on XML** (e.g. Define.xml file) – a potential for R to develop new tools.

CONSIDERATIONS FOR USING R IN REGULATED ENVIRONMENT

- **Security breaches**

In most cases security issues are not directly related to R but sometimes they do.

- **Industry standards**

- For computerized systems (possibility to use R in principle)
- For deliverables (electronic submissions to FDA)

- **Audits**

Sponsors, government agencies, independent parties.

RECAP

- **R usage varies across different stages of clinical research**
(inverse function of regulatory requirements)
- **R community has a great potential for developing new tools**
(interactive vs static reports, helpful for non-programmers, optimization of processes)
- **R is Okay for electronic submissions to FDA**
(need to overcome misconceptions and implement accordingly)
- **R can gain reputation by covering existing needs & challenges**
(e.g. better compliance with CDISC standards for electronic submissions)

Q&A SESSION