

Assuring Data Integrity and Data Quality in Sponsor Submissions

Paul Schuette
Scientific Computing Coordinator
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics
Paul.Schuette@fda.hhs.gov



Disclaimer

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

What does the FDA do?



FDA is responsible for protecting the public health by assuring the safety, efficacy and security of human and veterinary drugs, biological products, medical devices, our nation's food supply, cosmetics, and products that emit radiation.

See <https://www.fda.gov/AboutFDA/WhatWeDo/>

FDA regulated products comprise approximately 20% – 25% of consumer spending in the United States.

FDA Organization

The US FDA is organized primarily by centers:

- Center for Biologics Evaluation and Research (CBER)
- Center for Drug Evaluation and Research (CDER)
- Center for Devices and Radiological Health (CDRH)
- Center for Food Safety and Applied Nutrition (CFSAN)
- Center for Tobacco Products (CTP)
- Center for Veterinary Medicine (CVM)
- National Center for Toxicological Research (NCTR)
- Office of Regulatory Affairs

What does CDER do?

Promote public health by ensuring the availability of safe and effective drugs:

1. Conduct rigorous science-based premarket review to help ensure that drugs that will be marketed to the public are safe and effective
2. Identify and develop new scientific methods, models, and tools to improve the quality, safety, predictability and efficiency of new drug development
3. Promote patient and health professional awareness of drug benefits and risks through effective communication of drug information

CDER Responsibilities, post market



Protect public health by promoting the safe use of marketed drugs

1. Conduct post-market surveillance to enable early detection of new safety signals
2. Conduct rigorous studies to understand emerging drug safety signals and effectively manage those signals
3. Promote patient and health professional awareness of drug risks and safe use
4. Oversee drug promotion and marketing to help ensure that marketed drug labeling and advertising are truthful and not misleading

CDER Responsibilities, continued

Protect public health by ensuring the quality and integrity of marketed drug products

1. Secure the global supply chain to help ensure that drug integrity is maintained and that drugs are being manufactured and distributed to conform to established quality standards
2. Improve drug quality oversight capacity through expanded use of risk-based methods
3. Promote public and stakeholder awareness of drug quality and integrity issues through effective consumer communications

[FDA Center for Drug Evaluation and Research \(CDER\)
Strategic Plan 2013-2017](#)

Case Study 1

An inspector reported an incident in which several members of the same family were all enrolled in a clinical trial on a Friday evening. Other potentially questionable practices were found to have occurred at this site, which was the largest site in the trial. The sponsor withdrew the product from consideration after FDA expressed concerns regarding the site and the trial.

Case Study 2

Semler Research Private Limited. A 2015 inspection in Bangalore, India found significant instances of misconduct and violations of federal regulations, including the substitution and manipulation of study subject samples. As part of a coordinated action with WHO and EMA, letters were sent to sponsors notifying them of the need to repeat all bioavailability and bioequivalence studies performed by Semler,

<https://www.fda.gov/Drugs/DrugSafety/ucm495778.htm>.

This affected the U.S. President's Emergency Plan for AIDS Relief (PEPFAR).

Case Study 3

A clinical reviewer noticed that more adverse events were reported in a low dose arm than in a high dose arm in a recent clinical trial submission. When queried, it was discovered that the sponsor's CRO had incorrectly coded the treatment arms.

Case Study 4



A reviewer reported that a sponsor submitted a legacy data set in which the data quality was so poor as to be incapable of supporting a review. Among other findings, unique subject identifiers were not employed by the sponsor. Prior to receiving a complete response letter, the sponsor withdrew the product from consideration.

Data Quality Definition



Woodcock (NAS report, 1999) “...there is no consensus definition for ‘quality’ as it applies to data from clinical trials.”

ISO 8000: “Data quality is the degree to which data meets user requirements. ”

“The following principles of data quality underlie ISO 8000:

- a. **Data quality involves data being fit for purpose; i.e., the decision it is used in.**
- b. Data quality involves having the right data, in the right place, at the right time.
- c. Data quality involves meeting agreed customer data requirements.
- d. Data quality involves preventing the recurrence of data defects by improving processes to prevent repetition and eliminate waste.”

Data Integrity Definition

MHRA guidance :

“Data integrity is the degree to which data are complete, consistent, accurate, trustworthy, reliable and that these characteristics of the data are maintained throughout the data life cycle. The data should be collected and maintained in a secure manner, so that they are attributable, legible, contemporaneously recorded, original (or a true copy) and accurate. Assuring data integrity requires appropriate quality and risk management systems, including adherence to sound scientific principles and good documentation practices.”

<https://www.gov.uk/government/publications/guidance-on-gxp-data-integrity>

ALCOA: Attributable, Legible, Contemporaneously Recorded, Original (or true copy), and Accurate

Quality by Design

FDA has embraced the Quality by Design paradigm of Deming.

Some standards organizations:

- The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH)
- International Organization for Standardization (ISO)
- Clinical Data Interchange Standards Consortium (CDISC)
- Health Level Seven[®] (HL7)
- Over 100 others!

QbD, GxP

GxP = “Good x Practice”

Examples:

- Good Clinical Practice (GCP)
- Good Manufacturing Practice (GMP)
- Good Clinical Data Management Practice (GCDMP, Society for Clinical Data Management)
- Good Programming Practice (GPP, PhUSE)

E6(R2) Guidance

E6(R2) Good Clinical Practice: Integrated Addendum to ICH E6(R1) Guidance for Industry

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

March 2018
Procedural

OMB Control No. 0910-0843 Expiration Date 09/30/2020
See additional PRA statement in section 9 of this guidance.

New (March 2018) release

Specifies roles and responsibilities of:

- IRB/IEC
- Investigator
- Sponsor/CRO

E6 (R2) Section 4 Investigator



- “4.1.4 The investigator/institution should permit monitoring and auditing by the sponsor, and inspection by the appropriate regulatory authority(ies).”
- “4.9.7 Upon request of the monitor, auditor, IRB/IEC, or regulatory authority, the investigator/institution should make available for direct access all requested trial-related records.”
- Sponsors may need to update processes and procedures to ensure compliance with E6(R2) and General Data Protection Regulation (GDPR).

E6(R2), Section 5 Sponsor

Some subsections:

5.0 Quality Management

5.1 Quality Assurance and Quality Control

5.2 Contract Research Organization (CRO)

5.4 Trial Design

5.5 Trial Management, Data Handling, and
Recordkeeping

5.18 Monitoring

5.19 Audit

5.20 Noncompliance

E6(R2), Section 5.1

“5.1 Quality Assurance and Quality Control

- 5.1.1 The sponsor is responsible for implementing and maintaining quality assurance and quality control systems with written SOPs to ensure that trials are conducted and data are generated, documented (recorded), and reported in compliance with the protocol, GCP, and the applicable regulatory requirement(s).
- 5.1.2 The sponsor is responsible for securing agreement from all involved parties to ensure direct access (see section 1.21) to all trial-related sites, source data/documents, and reports for the purpose of monitoring and auditing by the sponsor, and inspection by domestic and foreign regulatory authorities.
- 5.1.3 Quality control should be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly.”

E6(R2), 5.18 Addendum



“On-site monitoring is performed at the sites at which the clinical trial is being conducted. Centralized monitoring is a remote evaluation of accumulating data, performed in a timely manner, supported by appropriately qualified and trained persons (e.g., data managers, biostatisticians).”

E6(R2), 5.18 Addendum, CSM

“Review that may include statistical analyses of accumulating data from centralized monitoring can be used to:

- a) Identify missing data, inconsistent data, data outliers, unexpected lack of variability and protocol deviations
- b) Examine data trends such as the range, consistency, and variability of data within and across sites
- c) Evaluate for systematic or significant errors in data collection and reporting at a site or across sites; or potential data manipulation or data integrity problems
- d) Analyze site characteristics and performance metrics
- e) Select sites and/or processes for targeted on-site monitoring.”

CDER Efforts

Inspectors from CDER include:

- Office of Scientific Investigations (OSI). Responsibilities include nonclinical and clinical drug product studies, bioequivalence studies, human subject protections in clinical drug product studies, post-market Adverse Drug Experience (PADE), Risk Evaluation and Mitigation Strategies (REMS), and, Postmarketing Requirements (PMR), and Safety Labeling.
- Office of Study Integrity and Surveillance (OSIS). Areas include pharmacokinetic, bioavailability/bioequivalence (BA/BE), Good Laboratory Practice (GLP), and Animal Rule (AR) studies.

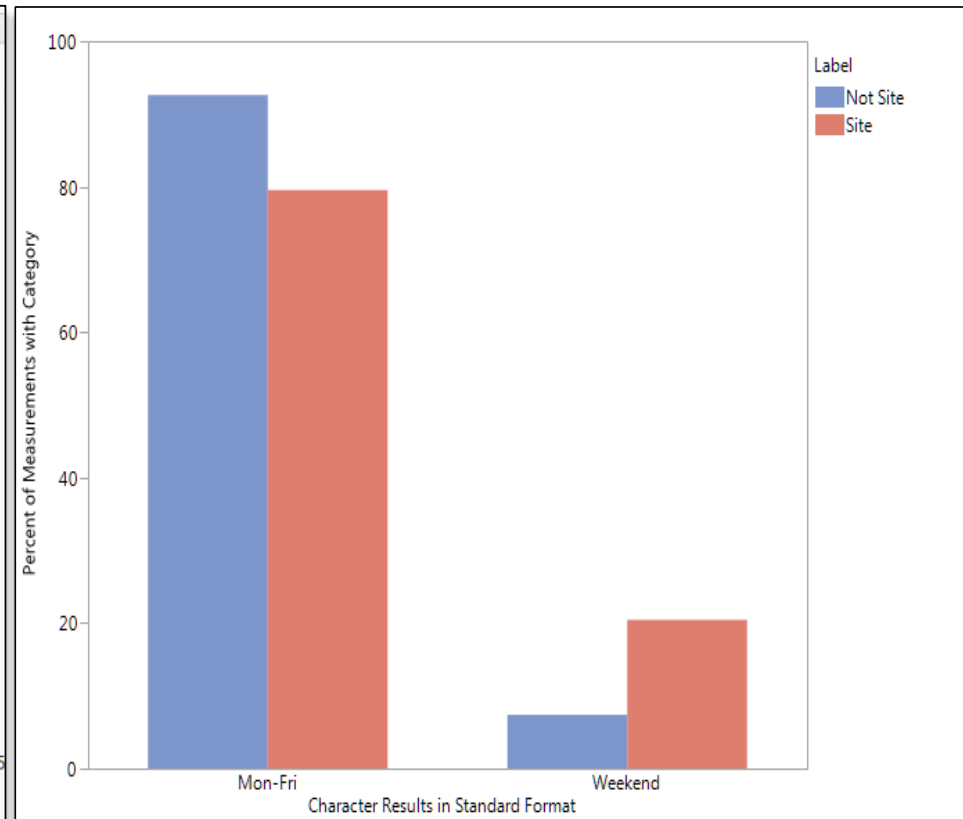
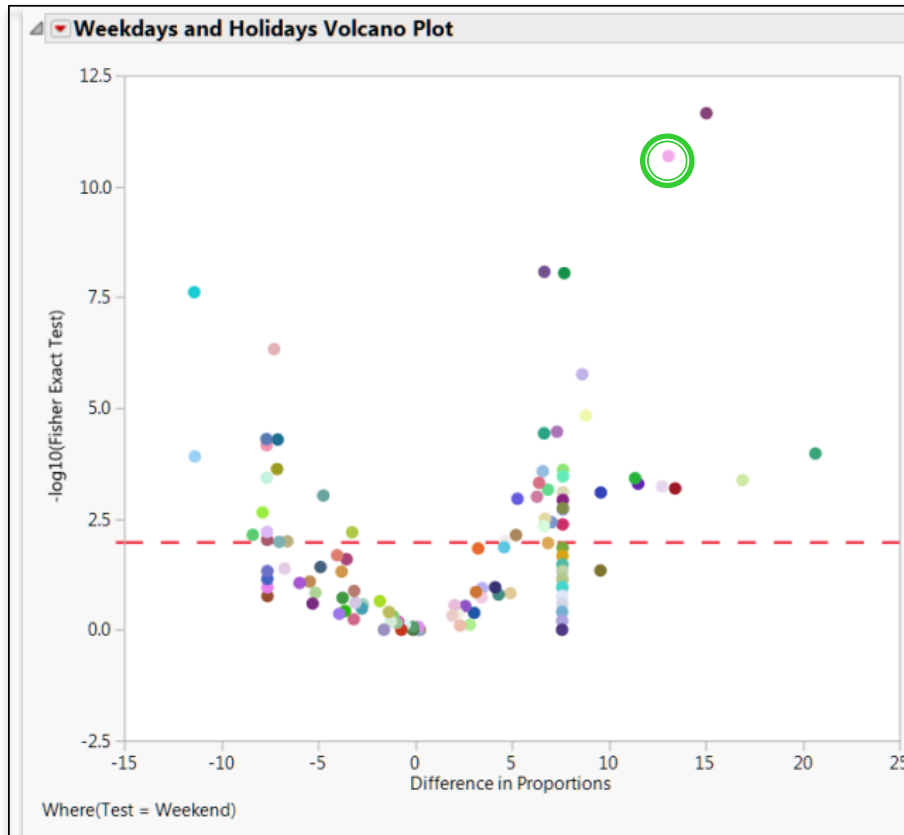
Software Tools

- Clinical Investigator Site Selection Tool (CISST). CISST is an Excel-based tool that analyzes site-level clinical trial data by study and provides a risk-rated prioritized list of clinical investigator sites for inspection.
- CISST Assist. An OCS pilot for using JMP Clinical in tandem with CISST.
- Core DataFitness Assessments (CoreDF). CoreDF summarize a list of common data quality errors in SDTM

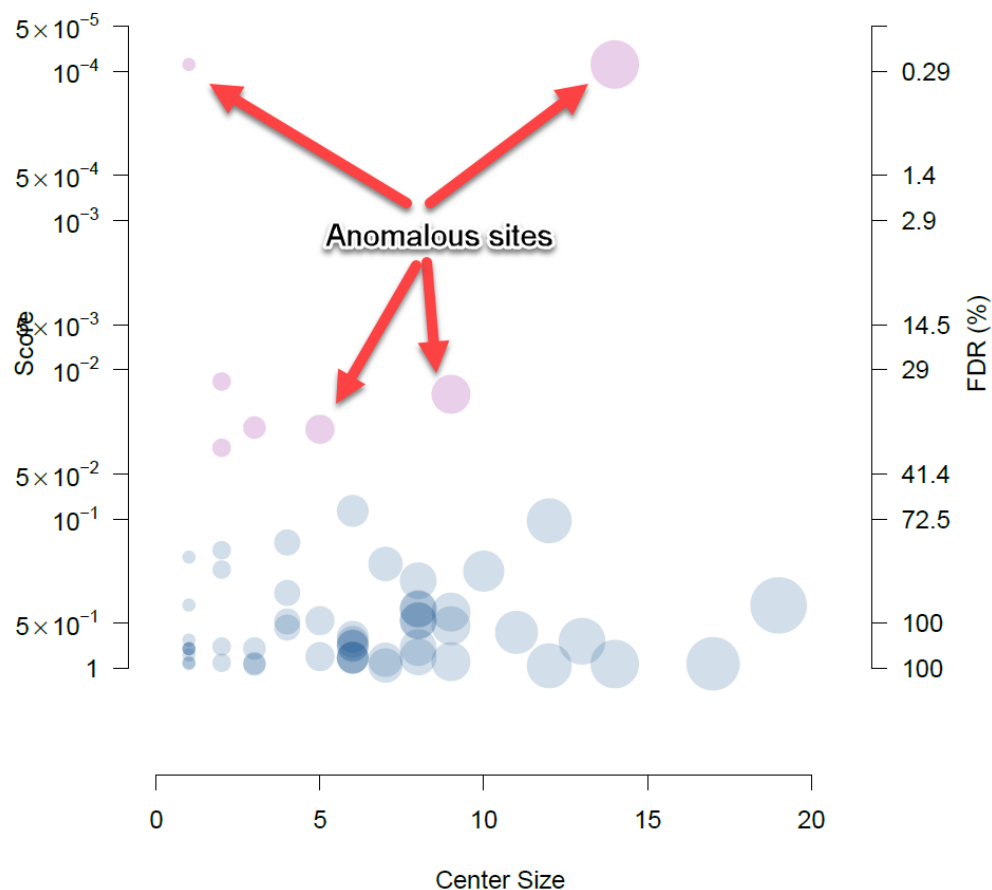
More Software Tools

- CluePoints Cooperative Research and Development Agreement (CRADA). The FDA and CluePoints have a three year CRADA in place to evaluate and further develop CluePoints's SMART™ engine for Risk Based and Central Statistical Monitoring. Using subject level data, the SMART engine applies a battery of statistical tests to variables across multiple data sets in a clinical trial and derives a data inconsistency score. Inspections and/or sensitivity analyses can then be conducted on potentially suspect sites.
- SAGE. A Python based tool for OSIS inspectors.
- Data Anomaly Detection Tool. A R Shiny app to explore and detect excessively similar PK/PD profiles in ANDAs for both reviewers and inspectors.
- Full disclosure: Paul Schuette is the PI for the CluePoints CRADA and is a consultant for the data anomaly detection tool.

CISST Assist Example, Proportion of Site Visits on a Weekend



Example of CRADA software output



Data from an approved product

- Most identified anomalous sites (magenta) were outside of US
- Treatment affect observed primarily outside of US
- Sponsor's audit revealed problems at the two largest anomalous sites
- CISST data not available for this trial

Inspections and Procedures

FDA inspectors (OSI and/or OSIS) may visit a site that has been identified as potentially problematic. Inspection outcomes:

- No Action Indicated (NAI)
- Voluntary Action Indicated (VAI)
- Official Action Indicated (OAI)

Rates: NAI > VAI >> OAI

“Fraud is rare, sloppiness is endemic.” (Marc Buyse)

Sensitivity Analyses: Sites identified as problematic may be excluded from statistical analyses.

Challenges

Progress has been made:

- Data standardization, CDISC requirements
- NAI rates have increased since 1999 NAS/IOM report

Challenges:

- Single Site ANDA Trials (Semler)
- Heterogeneity in multiregional trials
- Data Processing and Management
- Poor programming practices and computational reproducibility issues
- Premature unblinding
- Poor trial design and conduct
- Patient conduct

Conclusions

- Sponsors have a responsibility to assure data quality and data integrity in their submissions.
- Data quality and data integrity can be compromised in subtle ways.
- QbD, GxP can help ensure data quality and integrity.
- Software can identify data anomalies for follow up.
- FDA is concerned, the CDER Data Integrity Advisory Board has been recently formed to help address data quality and data integrity issues.

Key References

- Food and Drug Administration. (2016) Notification to Pharmaceutical Companies: Clinical and Bioanalytical Studies Conducted by Semler Research Are Unacceptable, <https://www.fda.gov/Drugs/DrugSafety/ucm495778.htm>.
- Davis, Nolan, Woodcock, Estabrook (editors). (1999) Assuring Data Quality and Validity in Clinical Trials for Regulatory Decision Making: Workshop Report, National Academy Press, <https://www.nap.edu/catalog/9623/assuring-data-quality-and-validity-in-clinical-trials-for-regulatory-decision-making>
- International Organization for Standardization. (2011) Data Quality-Part 1, Technical Specification ISO/TS 8000-1:2011(E), <https://www.iso.org/standard/50798.html>.
- Medicines & Healthcare products Regulatory Agency (2018) 'GXP' Data Integrity Guidance and Definitions, <https://www.gov.uk/government/publications/guidance-on-gxp-data-integrity> .
- Society for Clinical Data Management. (2013) Good Clinical Data Management Practices, <https://www.scdm.org/publications/gcdmp/>.
- Pharmaceutical Users Software Exchange. (2014) Guidance on Good Programming Practice, http://www.phusewiki.org/wiki/index.php?title=Good_Programming_Practice_Guidance.
- Food and Drug Administration. (2018) E6(R2) Good Clinical Practice: Integrated Addendum to ICH E6(R1) Guidance for Industry, <https://www.fda.gov/downloads/Drugs/Guidances/UCM464506.pdf>

Questions?



Statistical Software Clarifying Statement

“FDA does not require use of any specific software for statistical analyses, and statistical software is not explicitly discussed in Title 21 of the Code of Federal Regulations [e.g., in 21 CFR part 11]. However, the software package(s) used for statistical analyses should be fully documented in the submission, including version and build identification. ...

Sponsors are encouraged to consult with FDA review teams and especially with FDA statisticians regarding the choice and suitability of statistical software packages at an early stage in the product development process”

<https://www.fda.gov/downloads/ForIndustry/DataStandards/StudyDataStandards/UCM587506.pdf>