Statistical Science at FDA

Estelle Russek-Cohen
Director, Division of Biostatistics
Center for Biologics Evaluation and Research

1

Disclaimer

- Many of the thoughts are my own and may not necessarily reflect FDA perspectives.

- Acknowledgements

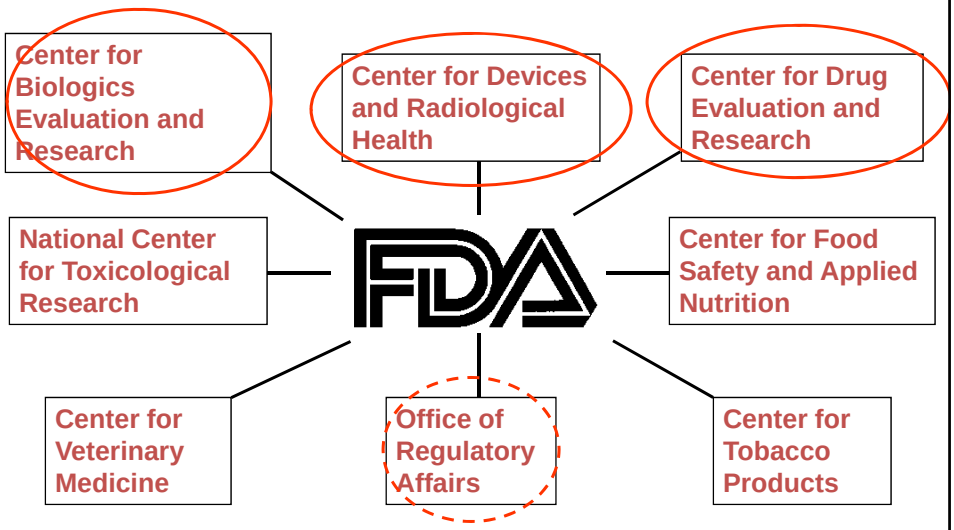
Thanks to Mike Nguyen for many of the slides on vaccine safety and to my colleagues on the tree-scan team.

2

Outline

- Overview of FDA
- The FDA Statistical Community
- Why two Adaptive Design Guidances?
- Estimands: Asking the right question of a clinical trial
- Personalized Medicine: Design Challenges
- Postmarket and big data
- Data and transparency
- Research funding, training and conclusions

3



4

FDA Statistical Community

- Over 300 ….statisticians in all the centers of FDA
- FDA Statistical Association:

  from all parts of FDA.
- Majority: reviews in support of decision-making at FDA. Includes evaluating study designs and study summaries submitted by “sponsors”.
- Many have a PhD; Many like being creative
- Research often motivated by reviews.

5

Statisticians at FDA

- Work in teams to review submissions
- Develop guidances (…work in teams)
- Active in many joint ventures with industry and academia (DIA, SCT and ASA working groups)
- Present talks at meetings: DIA-Stat Forum, JSM, ENAR, ASA Regulatory Workshop, etc.
- Mentor summer interns (ORISE)
- Senior statisticians serve on editorial boards of journals; are officers in statistical societies.

6

Medical Product evaluation:
Drugs/Biologics or Devices

- Most **diagnostics** are regulated as devices
- Most **therapeutics** are either drugs or biologics
- Development of diagnostics, therapeutic drugs/biologics and therapeutic devices are often different
- Typically different companies/sponsors are involved
- The statutes rather than the center typically drive regulation.
- When a diagnostic is used to decide who gets a new drug, may have more than one center involved.
- All 3 centers concerned with total product life cycle
- **Biggest challenge: Translation of acronyms….**

7

7

Guidance background

Levels of FDA policy

- Statutes are laws enacted by Congress
- Regulations are binding interpretations of the law
- Guidance documents are non-binding descriptions of FDA's current thinking on a topic
  - Often preceded by a public workshop on the topic
  - Initially released as draft for public comment
  - Comments received are discussed internally
  - Final versions are published that reflect comments rec'd

8

Guidance for Industry

Adaptive Design Clinical Trials for Drugs and Biologics

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact Robert O'Neill or Sue-Jane Wang at 301-796-1700, Marc Walton at 301-796-2600 (CDER), or the Office of Communication, Outreach and Development (CBER) at 800-835-4709 or 301-827-1800.

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)

February 2010
Clinical/Medical

- CDER/CBER Guidance
- Draft issued in 2010
- Public comments rec'd
- No discussion of devices
- Still a work in progress

9

Adaptive Designs for Medical Device Clinical Studies

Guidance for Industry and Food and Drug Administration Staff

Document issued on July 27, 2016.

The draft of this document was issued on May 18, 2015.

For questions regarding this document that relate to devices regulated by CDRH, contact Dr. Gerry Gray (CDRH) at 301-796-5750 or by e-mail at Gerry.Gray@fda.hhs.gov.

For questions regarding this document that relate to devices regulated by CBER, contact the Office of Communication, Outreach and Development (CBER) at 1-800-835-4709 or 240-402-8010.

U.S. Department of Health and Human Services
Food and Drug Administration

Center for Devices and Radiological Health

Center for Biologics Evaluation and Research

www.fda.gov

10

Two Guidances

- Different types of products means different emphasis; difference in statutes.
- Drugs/Biologics: More emphasis on dose finding.

  Product development differs:

  Phase 1, 2 and 3
- Devices: they evolve often

  Greater history of Bayesian methods
- Majority of the studies are similar. Most are for therapeutic indications.
- Relatively few adaptive designs for diagnostics but companion diagnostics has its own guidances. (Will discuss)

11

Research: Best Practices for Simulating Clinical Trials

- Role of simulation to evaluate clinical trials including:

  Properties of adaptive and/or Bayesian adaptive designs.
- Type 1 error?: usually one sided .025
- Robustness to assumptions
- Bayesian methods:

  borrowing vs noninformative priors?
- Bayesian hierarchical models

12 12

Surveys of FDA Adaptive designs

- CBER

Lin, M., S. Lee, B. Zhen, J. Scott, A. Horne, G. Solomon, E. Russek-Cohen (2016) CBER's Experience with Adaptive Design Clinical Trials. Therapeutic Innovation and Regulatory Science 50:195-203

- CDRH

Yang et. al., 2016. "Adaptive Design Practice at the Center for Devices and Radiological Health (CDRH), January 2007 to May 2013" Therapeutic Innovation and Regulatory Science 1-8.

13

Estimands

- Randomized Clinical Trials….the gold standard when they work
- What to do when:
  - patients drop out,
  - patients use rescue medications,
  - patients stay in trial but no longer are on assigned therapy.

How will treatment benefit get captured?

14

Statisticians Often Focus on the Analysis

- Estimands should come first….
  Capture the right clinical question
  Be specific: what is treatment benefit
  Post-randomization events?
- Estimators come next
- We are revisiting terms like:
  intent to treat and per protocol

- Given this impacts most therapeutic areas, many research opportunities exist.

15

References

- ICH documents including concept paper ICH E9(R1): See www.ich.org
- NRC (2010) The prevention and treatment of missing data in clinical trials. National Academies Press.

  Look for 2012 articles in NEJM and SIM by the authors.
- Lavange, LM and T Permutt (2015) A regulatory perspective on missing data in the aftermath of the NRC Report. SIM 35:2853-2864
- Permutt (2015) A taxonomy of estimands for regulatory clinical trials with discontinuations. SIM 35:2865-2875
- Mehrotra D et al (2016) Seeking Harmony: Estimands and Sensitivity Analyses for Confirmatory Clinical Trials. Clinical Trials 13: 456-458

16

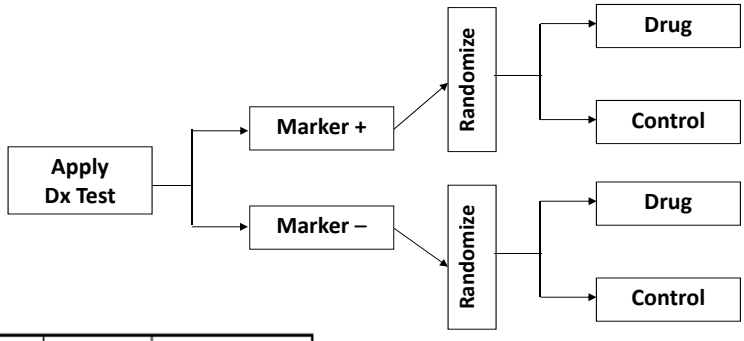
Personalized Medicine: Oncology

- A cancer Drug or Biologic will be approved by one center (drug/biologic)
- If a lab test is used to say "this drug is for you", lab test reviewed by another center (device)
- Study design challenges:

  One center: does the drug work in some?

  Other center: does the test help identify who?

17

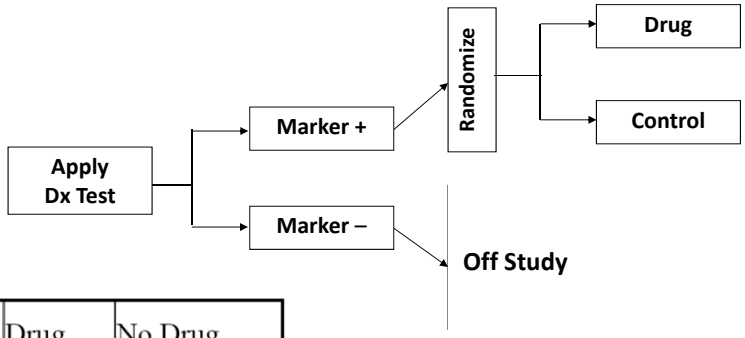


	Drug Resp	No Drug Resp
Test +	TP	FP
Test -	FN	TN

Can calculate performance estimators

- Sens = TP/[TP+FN]
- Spec = TN/[TN+FP]
- PPV = TP/[TP+FP]
- NPV = TN/[TN+FN]

18



	Drug Resp	No Drug Resp
Test +	TP	FP
Test -	-----	-----

- Can calculate
  - PPV

19

Which design is preferable?

- Which provides information on the ability of the biomarker to predict which patients will do well?
- Which is most efficient in evaluating the drug if the biomarker really selects the right patients?

20

References

- Pennello, G (2013) Analytical and clinical evaluation of biomarker assays: when are biomarkers ready for prime time. Clinical Trials (10): 666-676
- Li, M (2015) Statistical Methods for clinical validation of follow-on companion diagnostics via an external concordance study Stat Bioph Res (8): 355-363

21

References

- FDA (2016) Principles for Codevelopment of an In Vitro Companion Diagnostic Device with a Therapeutic Product

  Draft Guidance for Industry and FDA Staff
- FDA(2014) Guidance for Industry and FDA Staff: Qualification Process for Drug Development Tools

22

Postmarket and Big Data

- Seek info on safety that cannot be fully assessed in premarket
- Rarer adverse events
- Unanticipated events

  (incl: off-label use, drug interactions, etc)
- Very different studies than in premarket….also varies by data source.

23

Vaccines versus Drugs

- Vaccines:

  Fewer possible confounders with vaccines

  Limited exposures (eg 1, 2 or 3 shots)

  Often larger premarket studies
- Drugs: indications and duration can vary
- Larger premarket studies imply:

  Looking for very rare A.E.s in postmarket

24

24

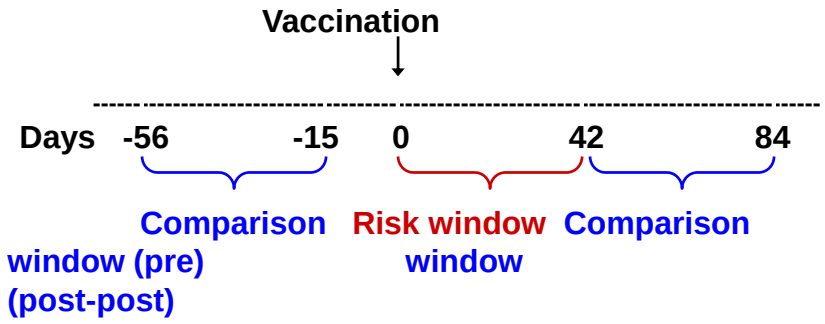
Self controlled case series

- Often used to study one product, one AE.
- Tutorial in Stat in Med (see references)
- Developed methods for vaccine safety
- Each subject serves as own control
- Efficient signal detection
- Commonly used in postmarket
- Doesn't formally address who is at risk

25

25

Self-Controls



26

Research

- Self Controlled methods control for confounders like sex, genetics, etc
- Concern over changes in very young children (time-varying confounders)
- Li,L., M. Kulldorff, E. Russek-Cohen, A. Tse Kawai, Wei Hua (2015) Quantifying the impact of time-varying baseline risk adjustment in the self-controlled risk interval design. Pharmacoepidemiology and Drug Safety 24:1304-1312

27

U.S. FOOD & DRUG ADMINISTRATION

Vaccine Adverse Event Reporting (VAERS)

Passive Surveillance

Events with no denominators

28

Passive Surveillance: VAERS

- Voluntary reports
- Patients, physicians, others
- Most fields are publicly available
- Case of no denominators…

  how many are exposed to product?
- CBER review: Division of Epidemiology
- Use Empirica Signal Detection Software

29 29

Passive Surveillance: AERS and VAERS

- **STRENGTHS:**
  - Open-ended for hypothesis generation
  - Potential detection of new or rare adverse events
  - Timeliness
  - Geographic diversity
  - Capability to monitor production lots
- **LIMITATIONS:**
  - Missing and inaccurate data
  - Under-reporting/Stimulated reporting
  - Absence of controls and denominators
  - Inability to assess causation
  - Low likelihood of detection for long latency events

30 30

CBER research initiative

- Text mining of narratives in VAERS
- 2 stage process:
- using natural language processing to extract features from text
- use supervised learning methods to develop classification rule.

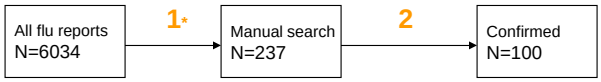
Can evaluation of narratives improve yield rate of anaphylaxis?

31

31

Review by Medical Officers

- Manual search and review of case reports for H1N1 anaphylaxis (10/12/2009-06/30/2010).



- Important to automate:
  - the whole process, but step 2 requires MR (<=>pdf files) review.
  - at least step 1 and provide MOs with the low number of reports for further review in step 2.

32

32

Training set: Classification Results

		MOs' review		
		Pos	Neg	Totals
Text Miner	Pos	183	352	535
	Neg	54	5445	5499
Totals		237	5797	6034

Sensitivity: 77.2%
Specificity: 94.0%

PPV: 34.2%
NPV: 99.0%

33

Independent validation

Independent validation of algorithm, truth determined by manual review (N=689):

PPV=30% NPV=99%

PPV hurt by low prevalence, but NPV suggests text miner can be used to enrich the dataset.

34

34

Other research in passive surveillance

- Papers by DuMouchel et al on Empirical Bayes methods
- Numerous papers by Ram Tiwari et al (CDER)
  - Using likelihood methods
  - Extensions considering covariate info (CBER)
- Detecting association rule mining for symptom clusters (CBER)

35

Vaccine Safety DataLink
CDC and FDA

36

CDC Vaccine Safety Datalink (1991)

- Eight geographically diverse health maintenance organizations that participate in a large linked database representing approximately 3% of U.S. population
- Surveillance and "Hypothesis testing" studies can be conducted
  - Vaccination (exposure)
  - Outpatient, emergency department, hospital and laboratory coding data (health outcomes)
  - Demographic variables (confounders)
  - Accessible medical chart review

37 37

VSD Rapid Cycle Analysis

- Method motivated by Wald SPRT:

  Wald: Simple vs Simple Hypotheses
- Near continuous monitoring (weekly)
- Extension to composite alternative
- Two variants:

  Poisson: #events vs expected counts

  Binomial: event rates exposed vs not exp

38 38

Max SPRT
Kulldorff et al (2011)

- Likelihood ratio statistic
  Poisson or binomial
- Length of surveillance fixed (e.g. 2 yrs)
- Time is expected counts not calendar time
- Rapid detection important
- A number of signals were identified, most are false alarms.

39

PRISM Basics

- Mini-Sentinel Pilot & Sentinel program dedicated to **vaccine** safety (see www.sentinel.org)
- Claims based system with data from 4 national health plans
  - Aetna, HealthCore (Wellpoint), Humana, Optum (United Healthcare)
  - Data linked to 8 vaccine registries in USA
- Access to medical records and pharmacy data

Pharmacoepidemiol Drug Saf. 2012 Jan;21 Suppl 1:291-7.

40

Prespecified event: Some methods

1. **Self-controlled design**
   - Useful for single or short-term exposures or when no independent comparator group is available
   - When between-person confounding is large but within-person confounding is modest
2. **Exposure match cohort**
   - Uses propensity or disease risk scores in fixed or variable ratio
   - Provides flexible choices of effect measures, multiple endpoints and broad range of alerting rules
3. **Full cohort design with regression**
   - Permits a high degree of analytic flexibility (e.g., the ability to simultaneously evaluate interactions, multiple comparison groups, and subgroups)

41 41

Data Mining Development

- Test whether it is possible to detect adverse events without pre-specifying them a priori
- Develop statistical approach to simultaneously evaluate hundreds of different adverse events
  - Advantage: detect unexpected adverse events
  - Disadvantages: not possible to adjust for all possible confounders, as they vary by disease outcomes

    Finding optimum risk window for all events is hard

    Hierarchy of events imperfect.

  Pilot phase: emphasis on detecting known signals.

42 42

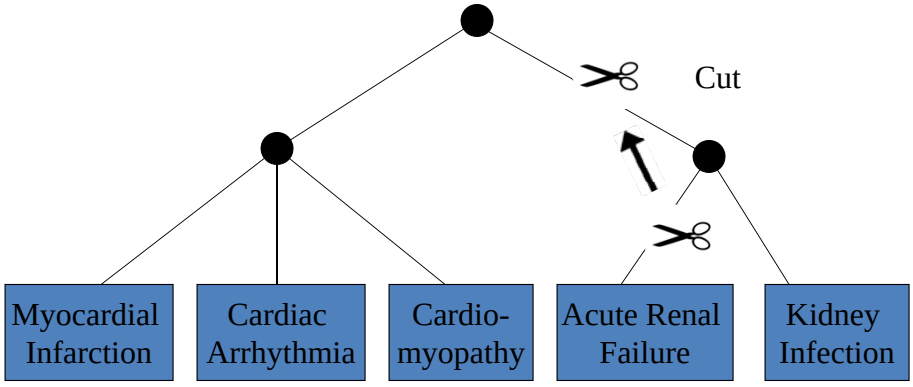
Mining: 3 Methods Being Evaluated

- Project led by Martin Kulldorff
  - DuMouchel's Gamma Poisson Shrinker
  - Tree-based scan statistic with population based controls
  - Tree-based scan statistic with self-controls
- Basics of Tree Scan algorithm
  - Use a hierarchical tree
  - Evaluate cuts on the tree (assess observed vs. expected at each leaf)
  - Control for multiple testing

Kulldorff M, Fang Z, Walsh S. A tree-based scan statistic for database disease surveillance. Biometrics, 2003,59:323-331.

43

Example of a Small Tree



44

Tree-based SCAN Statistic

1. Scan the tree by considering all possible cuts on any branch
2. For each cut, calculate the likelihood
3. Denote the cut with the maximum likelihood as the most likely cut (cluster)
4. Generate 9999 Monte Carlo replications under H_0.
5. Compare the most likely cut from the real data set with the most likely cuts from the random data sets
6. If the rank of the most likely cut from the real data set is R, then the p-value for that cut is R/(9999+1).

45 45

Tree-based SCAN Statistic

1. Scan the tree by considering all possible cuts on any branch
2. For each cut, calculate the likelihood

Helps answer, "Has FDA observed any new safety issues?" without pre-specifying a particular outcome

4. Generate 9999 Monte Carlo replications under H_0.
5. Compare the most likely cut from the real data set with the most likely cuts from the random data sets
6. If the rank of the most likely cut from the real data set is R, then the p-value for that cut is R/(9999+1).

46 46

References

- Text mining for the Vaccine Adverse Event Reporting System: medical text classification using informative feature selection
  Botsis T, Nguyen MD, Woo EJ, Markatou M, Ball R (on line June 2011)
- Tutorial in Biostatistics: The self controlled case series method SIM 2006 25: 1768-1797
  Whitaker, HJ, Farrington CP, Spiessens B, Musonda, P.
- Post-Licensure Rapid Immunization Safety Monitoring AJE (2012)
  Yih WK, Lee G, Lieu T, Ball R, Kulldorff M, et al
- Bayesian Data Mining in Large Frequency Tables with an Application to FDA Spontaneous Reporting System. Am. Stat. 1999 53:177-190
  DuMouchel, W. Often used at FDA with passive surveillance data.
- M. Kulldorff et al. Tree Scan 1.3 Software with User Guide
  http://www.treescan.org

47

47

Statisticians want data:

Transparency

- Clinical Trial data submitted to FDA:
  Belongs to the sponsor of the trial
- Some companies participating in programs to make de-identified data available (GSK;YODA)
- Possible concerns over re-analysis
  New insights or data dredging?
- We need to find ways to overcome this.

48

Most sources of big data at FDA

- Sentinel and federal agencies:
  - Data is behind firewalls
  - Patient privacy is protected
- VSD, and VAERs have de-identified public datasets.

49

Research Funding

- CERSI
  - Centers for Excellence in Regulatory Science & Innovation
  - Includes UM, Stanford, Hopkins, Georgetown
- BAA
  - Broad Agency Announcements
- Summer internships in CDER (ORISE)
- Some faculty have come on sabbatical leave

50

Training for a Career at FDA

- Foundation in statistical theory essential
- Course in clinical trials helpful
- Written and oral communications are important
- Critical thinking skills essential too….get practical experience while in school!
- Useful skills: computation and simulations, bayesian methods, data mining, ….
- **Expect to keep learning to survive at FDA**

51

Conclusions

- FDA Statisticians have made many contributions to statistics, to the profession and to the mission of the agency
- New challenges keep coming
- Several UMBC and UMCP graduates at FDA…

  …we would also welcome Stanford graduates
- We think it is a great place to have a career.

52


FDA
U.S. FOOD & DRUG
ADMINISTRATION