What Matters in Clinical Trial Quality

David L DeMets, PhD

Department of Biostatistics &
Medical Informatics
University of Wisconsin-Madison

DeMets COI

- I consult with
  - NIH
  - FDA
  - IOM
  - Industry
- I serve on several trial DMCs currently
- I hold no stock in biopharm/device industry
- I am not selling any medical products

Clinical Trial Quality

- Experienced leadership
- Committed sponsor
- Adequate network of clinical sites
  - Experienced
  - Proven ability to recruit subjects
  - Committed to scientific question
- Data management system
- Statistical team

Clinical Trial Quality (2)

- Scientific basis for primary hypothesis
  - Not always adequate
  - Preliminary studies may be too small
- Well defined hypothesis
- Correct Trial Design for the Hypothesis
- Randomization for Phase III
- Key outcomes that are
  - Reproducible
  - Sensitive to proposed intervention
  - Clinically Relevant

Clinical Trial Quality (3)

- Adequately sized / powered
- Subjects relevant to question being asked – not likely a random sample
- Blinded intervention or at least blinded evaluation
- Good subject compliance to the protocol
- High ascertainment in key outcomes
  - Primary, secondary
  - SAEs
- Follow ITT Principle for analysis

Data Quality

- Data quality where it matters
- Some amount of error can be tolerated
  - Random error
  - Increasing sample size can average it out
  - Will dilute intervention effect
- Other error cannot be tolerated
  - Systematic error or bias
  - Missingness
    - Missing data usually not missing at random
    - Censoring can be informative

Data Quality Issues

- Not all data are created equal
- Some data more critical than others
- Not cost effective to get “perfect’ data for all variables measured
- Need to know: Real patients, real disease, real treatment and real outcomes
- 80/20 Rule for data quality
- Two natural experiments

Is 100% CRF validation cost effective?

- Two natural experiments
  - GUSTO-I trial of tpa in cardiac patients
    - NEJM, 1993
  - Breast cancer trial
    - NEJM, 1995
- Some errors detected but no influence or even noticeable affect on analysis

Gusto-I
(NEJM, 1993)

- Industry sponsored trial totally managed by academic network (Duke)
  - Conducted with a central statistical & data coordinating center
  - Standardized case report forms
  - Central statistical quality control utilized
  - Positive results for mortality & morbidity
- Later, sponsor did 100% case report form audit for regulatory submission
- Some errors found, but no bias & no effect on statistical analysis or results

Breast Cancer Trial
(NSABP, NEJM, 1995)

- NIH sponsored trial (NSABP)
- Compared radical vs lumpectomy surgery
- Statistical center discovered data anomalies in biopsy forms at one site
- A few dates altered to avoid repeat biopsy
- Public discovery of this led to massive investigation

Breast Cancer Fraud

- With verification of altered dates, all data from selected other sites audited intensively
- Some errors in data found
- No real changes in analysis results
- Fraud often discovered by statistical anomalies or irregularities in the data
- Central data auditing can be very effective, at less cost

Most Critical Data

- High accuracy desired; these include
  - Eligibility Criteria
  - Primary Outcome
  - Key Secondary Outcomes
  - SAEs
  - Selected AEs
  - Lost to FU information
  - Lab data that are “deal breakers”
    - Eg QT interval, LFTs,..

Data of Interest

- Accuracy desired but at a lesser degree
  - History and physical data
  - Compliance
  - Larger but still limited list of AEs
  - Lab data related to mechanism
  - Etc

Data of Lesser Interest

- Some data of interest but difficult to utilize in analysis of outcome data
- Tend to be expansive, hard to QC
- These include:
  - Detailed concomitant medications
  - Medra level type AE coding

Standard Coding Systems

- Several adverse event coding systems
- Often organized by body systems
- Subcategorized into events as reported by investigator/patient – passive system
- Safety committees often review these in tabular form by treatment arms
- (AE listings not helpful after awhile)

- UNCODED

Serious Adverse Events	N Subjects (Events)			Percent of Subjects			
Preferred Term	A	B	All	A	B	All	pA.B
OVERALL	16 (17)	22 (22)	38 (39)	0.8	1.1	1.0	0.337
no term given	16 (17)	22 (22)	38 (39)	0.8	1.1	1.0	0.337

BODY AS A WHOLE - GENERAL DISORDERS

Serious Adverse Events	N Subjects (Events)			Percent of Subjects			
Preferred Term	A	B	All	A	B	All	pA.B
OVERALL	118 (384)	167 (573)	285 (957)	5.9	8.3	7.1	0.003
accident and/or injury	8 (9)	9 (9)	17 (18)	0.4	0.4	0.4	0.817
alcohol intoxication	0 (0)	4 (4)	4 (4)	0.0	0.2	0.1	0.046
alcoholism intox	0 (0)	1 (1)	1 (1)	0.0	0.0	0.0	0.319
animal bite/scratch	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
ankle oedema	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
anticoagulant effect increased	0 (0)	1 (1)	1 (1)	0.0	0.0	0.0	0.319
ascites	1 (1)	2 (2)	3 (3)	0.1	0.1	0.1	0.567
asthenia	0 (0)	1 (1)	1 (1)	0.0	0.0	0.0	0.319
back pain	2 (2)	0 (0)	2 (2)	0.1	0.0	0.1	0.156
back pain (lumbago)	2 (2)	1 (1)	3 (3)	0.1	0.0	0.1	0.560
chest discomfort	0 (0)	3 (3)	3 (3)	0.0	0.1	0.1	0.084
chest pain	24 (27)	35 (39)	59 (66)	1.2	1.7	1.5	0.155
chest pain (angina like)	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
collapse nos	1 (1)	2 (2)	3 (3)	0.1	0.1	0.1	0.567
concussion brain	4 (4)	0 (0)	4 (4)	0.2	0.0	0.1	0.045
death	11 (11)	19 (19)	30 (30)	0.6	0.9	0.8	0.147
digitalis intoxication	5 (5)	5 (5)	10 (10)	0.3	0.2	0.3	0.993
dislocated shoulder	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
dislocation finger	0 (0)	1 (1)	1 (1)	0.0	0.0	0.0	0.319
esr increased	1 (1)	1 (1)	2 (2)	0.1	0.0	0.1	0.997
extreme fatigue	1 (1)	1 (1)	2 (2)	0.1	0.0	0.1	0.997
fall, external reason	0 (0)	1 (1)	1 (1)	0.0	0.0	0.0	0.319
feeling unwell	1 (1)	1 (1)	2 (2)	0.1	0.0	0.1	0.997
fever	5 (5)	3 (3)	8 (8)	0.3	0.1	0.2	0.474
flank pain	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
flu symptoms	0 (0)	1 (1)	1 (1)	0.0	0.0	0.0	0.319
injury knee	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
malaise	0 (0)	2 (3)	2 (3)	0.0	0.1	0.1	0.158
oedema	0 (0)	2 (2)	2 (2)	0.0	0.1	0.1	0.158
oedema leg(s)	1 (1)	1 (1)	2 (2)	0.1	0.0	0.1	0.997
oedema peripheral	0 (0)	1 (1)	1 (1)	0.0	0.0	0.0	0.319
overdose	1 (1)	1 (1)	2 (2)	0.1	0.0	0.1	0.997
pacemaker malformation	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
pacemaker malfunction	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
pain arm	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
pain groin	0 (0)	1 (1)	1 (1)	0.0	0.0	0.0	0.319
pain shoulder	1 (1)	1 (1)	2 (2)	0.1	0.0	0.1	0.997
pain/ache in leg(s)	3 (3)	1 (1)	4 (4)	0.2	0.0	0.1	0.314
postop abscess	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
postop wound haematoma	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
postop wound infection	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
postoperative complications	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
prosthesis disorder	1 (1)	2 (3)	3 (4)	0.1	0.1	0.1	0.567
sciatica	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
strained neck	1 (1)	0 (0)	1 (1)	0.1	0.0	0.0	0.316
strength loss of	0 (0)	1 (1)	1 (1)	0.0	0.0	0.0	0.319
sudden cardiac death	0 (0)	1 (1)	1 (1)	0.0	0.0	0.0	0.319
sudden death	30 (30)	73 (74)	103 (104)	1.5	3.6	2.6	0.000
tiredness or fatigue	5 (5)	4 (4)	9 (9)	0.3	0.2	0.2	0.732
weakness	1 (1)	1 (1)	2 (2)	0.1	0.0	0.1	0.997

Standard Coding Systems

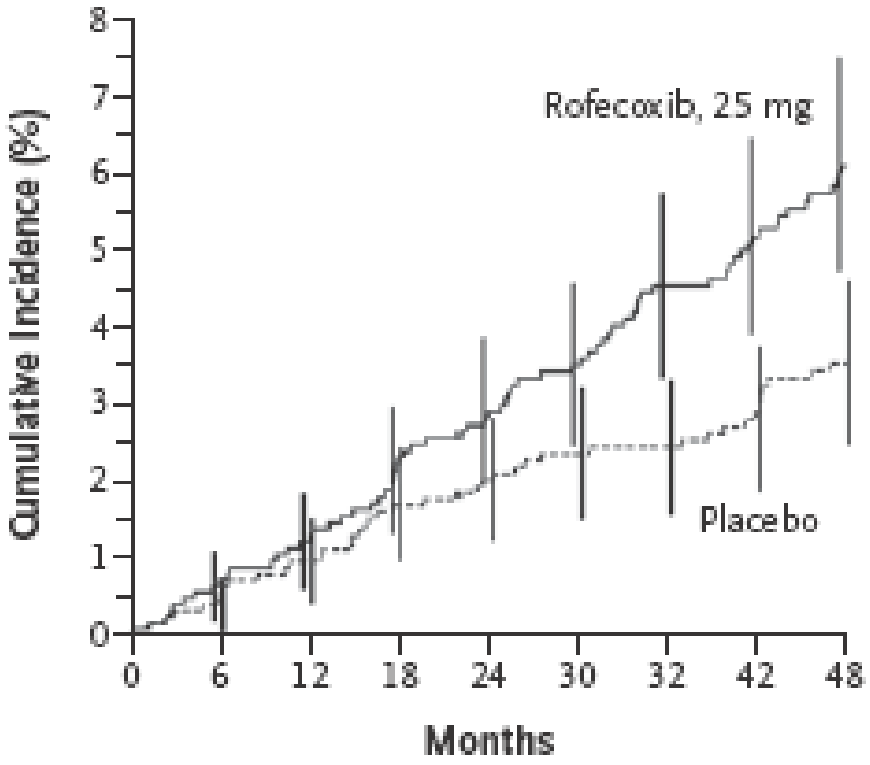
- Generally these tables not helpful
- At one level, too granular and tables filled with small number of events – difficult to interpret
- At higher level, too many critical and non critical events pooled together – also difficult to interpret
- DMCs often need to select critical events and create new safety variable(s), best not to be done post hoc
- Need to sharpen our safety focus on Adverse Events of Special Interest (AESIs)

Off Drug ≠ Off Study

- ITT requires inclusion of
  - All patients randomized
  - All events during follow up
- Exclusion of either patients or events can lead to bias
  - Direction is not always predictable
- If all events not captured, no way to tell if it makes a difference
- Censoring for withdrawal may be informative

APPROVE Trial

- References
  - NEJM 2005 Primary Paper
  - NEJM 2006 Editorials
  - Lancet 2008 Approve+1
- A trial of Vioxx (Rofecoxib) for colon cancer prevention
- 2005 Paper suggested an increase in CV events
- Debate over 18 month honeymoon



No. at Risk

Rofecoxib	1287	1221	1187	1152	1131	1117	1092	1032	989
Placebo	1300	1247	1224	1189	1173	1157	1133	1071	1027

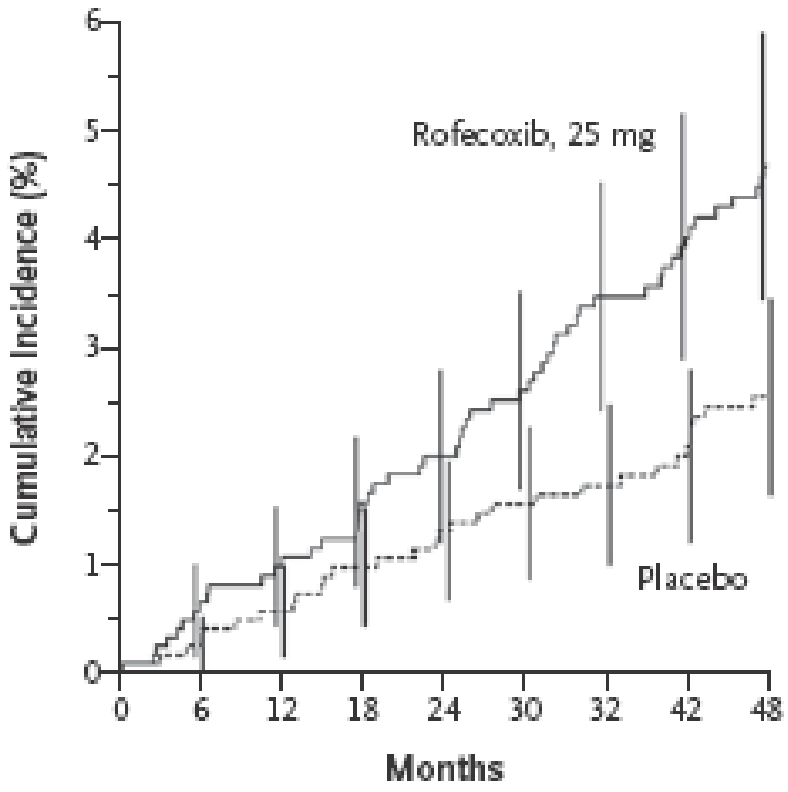
Figure 2. Kaplan–Meier Estimates of the Cumulative Incidence of Confirmed Thrombotic Cardiovascular Events in the Rofecoxib and Placebo Groups, According to the Intention-to-Treat Principle.

I bars represent 95 percent confidence intervals.

NEJM 2006; 355:203-05

Approve + 1
(Lancet, 2008)

- In initial design, patients who went off drug were not followed after 14 days
- Pressures caused sponsor to conduct an additional year of follow-up on all patients randomized
- An independent analysis was conducted at UW on updated data
- Results did not confirm the 18 month honeymoon for CV risk



No. at Risk

Rofecoxib	1287	1220	1188	1158	1140	1125	1102	1042	1002
Placebo	1300	1249	1228	1196	1181	1165	1140	1079	1036

Figure 1. Kaplan–Meier Estimates of the Cumulative Incidence of Confirmed APTC Events in the Rofecoxib and Placebo Groups, According to the Intention-to-Treat Principle.

I bars represent 95 percent confidence intervals.

NEJM 2006; 355:203-05

Lessons on Censoring

- Censoring for any reason other than administrative may be informative
- Need to collect all follow-up data, regardless of adherence
- Must have all the data to assess whether censoring might have made a difference
- Too frequent of a problem (7, 14, 30 day rules)
- Off drug ≠ off study !

New Challenges

- Omic data
  - Predictor of risk
  - Prognostic for response
- Enormous quantity of data
- What does quality mean?
- So far, omics data often not subjected to standard lab QC measures

Summary

- Need strong and experienced leadership
- Solid scientific basis for hypothesis
- Proper trial design
  - Appropriate outcomes
  - Adequate sample size
- Excellent compliance
- Complete and unbiased ascertainment
- ITT Analysis
- Prioritize QC for data