Global Hypothesis Testing in Adverse Effect Assessment

Zhibao Mi MD PhD and Joseph Collins Sc.D.

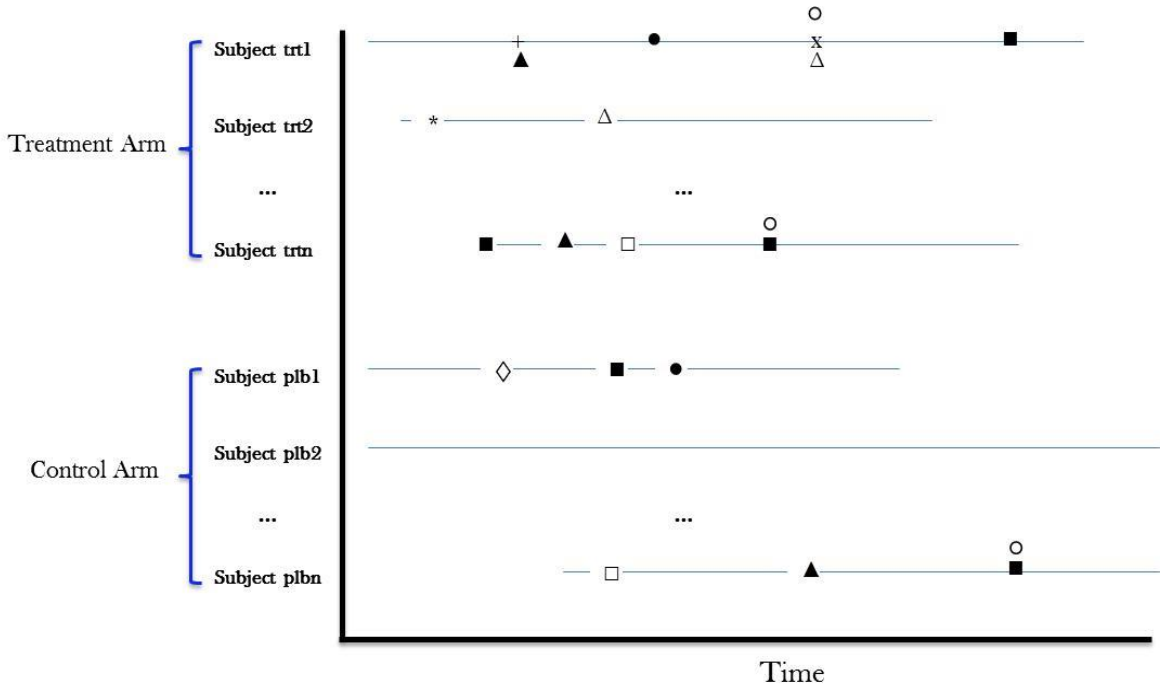
VA Cooperative Studies Program,
VA Maryland Health Care Systems, Perry Point, MD 21902

SCT 2014 - PHILADELPHIA

Motivation

- Clinical trials inference focused on efficacy endpoints
- Adverse effect analysis often lack using formal statistical methods
- Experts in the field demand rigorous adverse effect analysis
  - O'Neil (FDA, 1995, 2001, 2008)
  - Chuang-Stein (Pfizer, 1992, 1998)
  - Wittes (SCI, 1998, 2008, 2011)
  ...
- Our efforts to improve adverse effect analysis
  - Analytical approach (2011 JSM)
  - Simultaneous confidence interval (2012 SCT)
  - Lump and split strategy (2013 SCT)
  - Global hypothesis testing (2014 SCT)
- Global hypothesis testing : Comprehensive and sensitive

Adverse Events and Parameterization



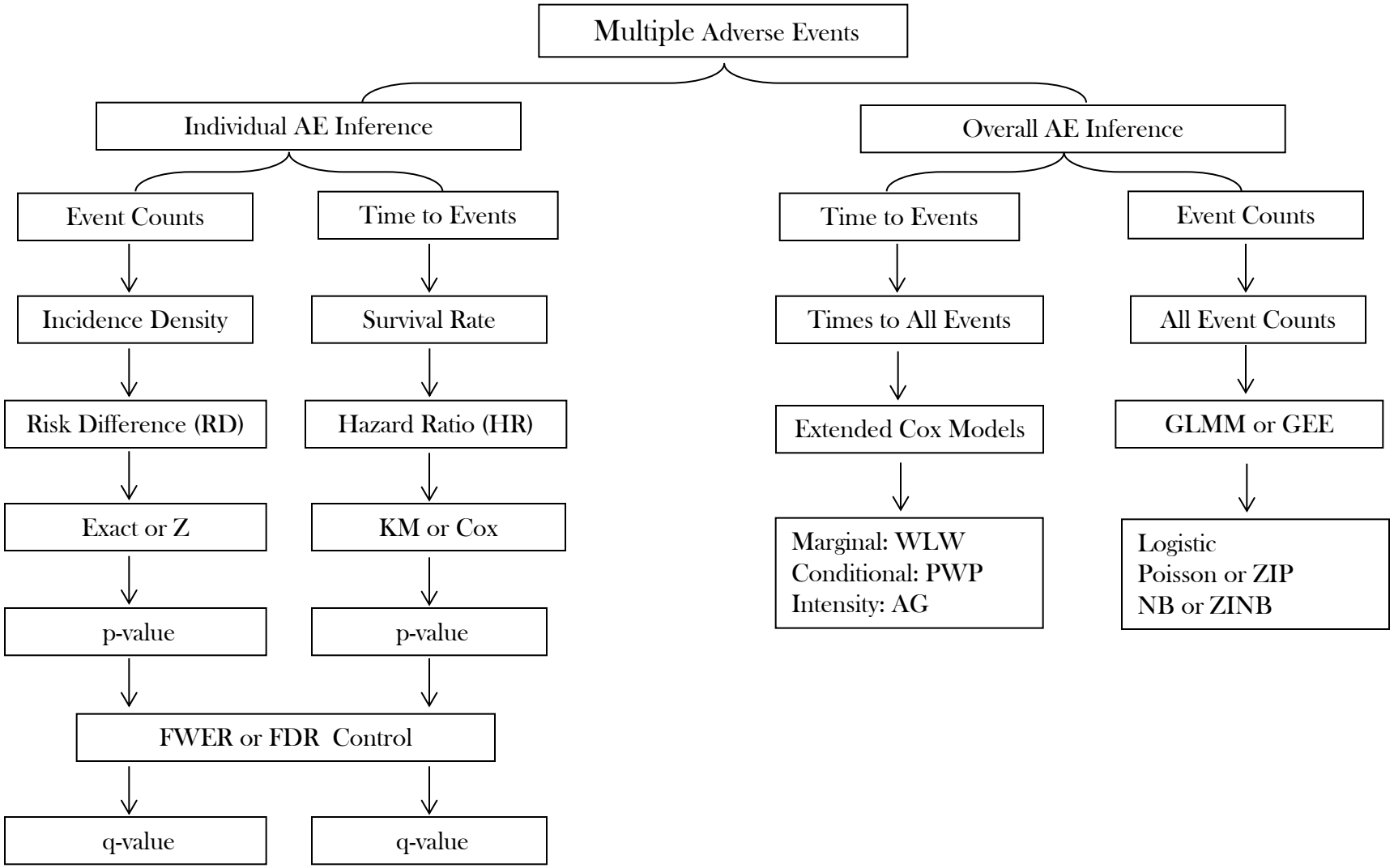
Data Parameterization

- Events: Multiple Counts Data (Correlated Cluster Data)
- Duration : Time-to- Multiple Events Data (Competing Risk Data)

Issues in Adverse Event Analysis

- Power Issue (Rare incidence, Not powered sample size)
- Multiplicity Issue (So many AE types)
- Correlation and Competing Risk Issue (Multiple adverse events)
- Sensitivity and Comprehensiveness (Global vs. individual)

Strategy for Multiple Adverse Event Analysis



Trial One

Trial –

- A placebo-controlled, double-blind, multi-center trial
- 212 subjects enrolled at 16 sites during 4 and half years
- 108 randomized to treatment group (trt) and 104 to placebo group (plb)

Data –

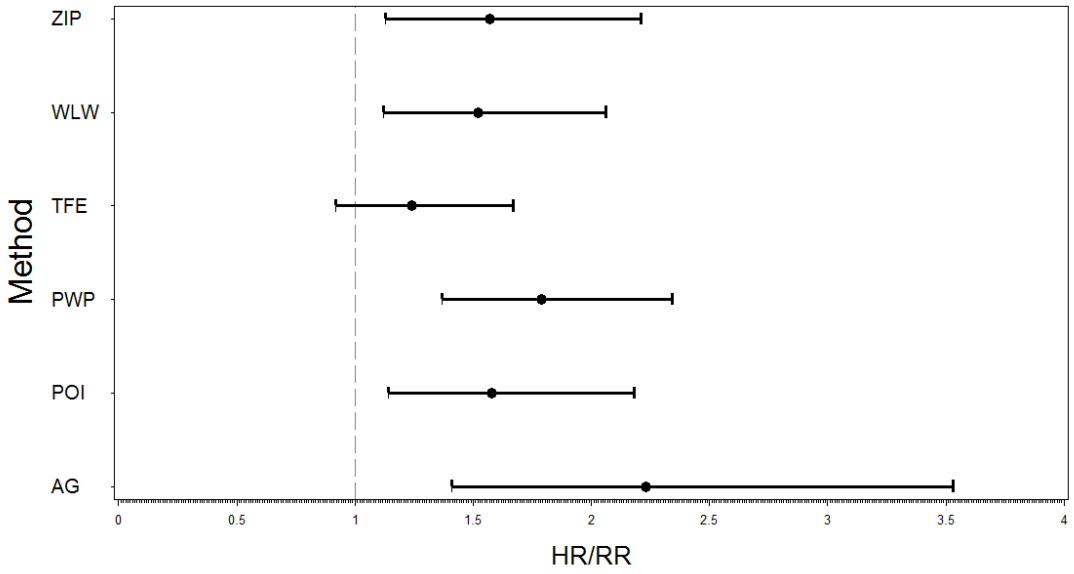
- 1348 AEs observed during the study
- 372 distinct AE symptoms based on preferred terms
- AEs multivariate binary variables
- Rare Events

Individual AE Analysis

AE	ID (trt)	ID (plb)	RD	p	q
AUTONOMIC NERVOUS SYSTEM IMBALANCE	0.1167	0.0043	0.1124	0.0000	0.0000
ALANINE AMINOTRANSFERASE INCREASED	0.0455	0.0000	0.0455	0.0000	0.0002
PYREXIA	0.0989	0.0297	0.0692	0.0000	0.0009
HEPATIC ENZYME INCREASED	0.0435	0.0021	0.0414	0.0000	0.0009
BODY TEMPERATURE INCREASED	0.0613	0.0170	0.0443	0.0003	0.0216
ASPARTATE AMINOTRANSFERASE INCREASED	0.0198	0.0000	0.0198	0.0014	NS
BLOOD POTASSIUM INCREASED	0.0138	0.0000	0.0138	0.0077	NS
BLOOD PRESSURE INCREASED	0.0138	0.0000	0.0138	0.0077	NS
HYPOTENSION	0.0237	0.0043	0.0195	0.0085	NS
PLATELET COUNT INCREASED	0.0119	0.0000	0.0119	0.0137	NS
DEBRIDEMENT	0.0158	0.0021	0.0137	0.0211	NS
HYPERKALAEMIA	0.0158	0.0021	0.0137	0.0211	NS
VOMITING	0.0119	0.0340	-0.0221	0.0218	NS
ANAEMIA	0.0237	0.0064	0.0174	0.0241	NS
HAEMOGLOBIN DECREASED	0.0178	0.0043	0.0136	0.0400	NS
BLOOD ALKALINE PHOSPHATASE INCREASED	0.0000	0.0085	-0.0085	0.0446	NS
OPEN WOUND	0.0040	0.0170	-0.0130	0.0476	NS
...	...	...	...	...	...

Global AE Analysis

Methods	HR / RR	95% C.I.	p
Time to Event Approach			
Time to First Event (TFE) Model	1.24	0.92 - 1.67	0.1519
AG Intensity Model	2.23	1.41 - 3.53	0.0006
WLW Marginal Model	1.52	1.12 - 2.06	0.0075
PWP Conditional Model	1.79	1.37 - 2.34	< 0.0001
Event Count Approach			
Poisson Model (POI)	1.58	1.14 - 2.18	0.0058
ZIP Model	1.57	1.13 - 2.21	0.0082



Trial Two

Trial –

- A placebo-controlled, double-blind, multi-center trial
- 68 subjects enrolled at 3 sites
- 35 randomized to treatment group (trt) and 33 to placebo group (plb)

Data –

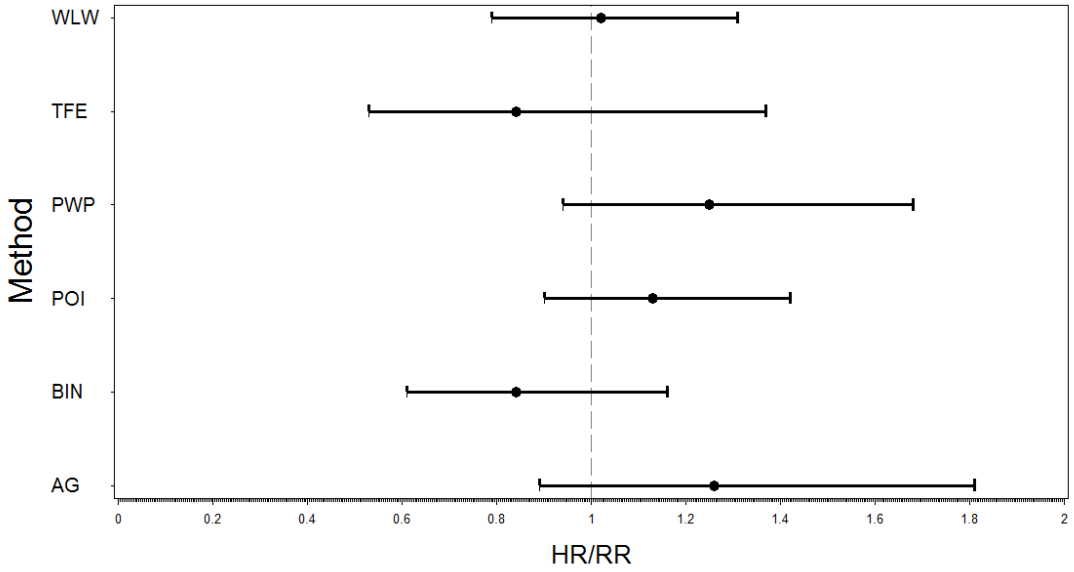
- 32 AEs collected during the study
- 30 distinct AE symptoms based on preferred terms
- AEs multivariate binary variables

Individual AE Analysis

AE	ID (trt)	ID (plb)	RD	p	q
HYPOTENSION	0.04369	0.00000	0.043689	0.00216	NS
VOMITING?EMESIS	0.03883	0.10596	-0.067125	0.01824	NS
RHINORRHEA	0.15534	0.25166	-0.096316	0.02649	NS
NAUSEA	0.06796	0.13907	-0.071112	0.03206	NS
DROWSINESS	0.08252	0.03311	0.049412	0.04012	NS
…	…	…	…	…	…

Global AE Analysis

Methods	HR / RR	95% C.I.	p
Time to Event Approach			
Time to First Event (TFE)	0.84	0.53 – 1.37	0.4959
AG Intensity Model	1.26	0.89 – 1.81	0.1967
WLW Marginal Model	1.02	0.79 – 1.31	0.8861
PWP Conditional Model	1.25	0.94 – 1.68	0.1262
Event Count Approach			
Logistic Regression Model (BIN)	0.84	0.61 – 1.16	0.2947
Poisson Model (POI)	1.13	0.90 – 1.42	0.2943



Summary

- AE analyses for both trials showed
  - Overall and individual testing results consistent
  - Duration and Event parameterized AE data yielded similar results
- Given the sample sizes for each trial, the testing powers
  - Trial one : 93.5% (POI GEE) Need 184 to yield 90% power (n = 212)
  - Trial two : 14.0% (POI GEE) Need 957 to yield 90% power (n = 68)
- AE assessment based on both overall and individual hypothesis tests
  - Robust assessment
  - Interpretation