Promoting Effective Drug Development Programs: Opportunities and Priorities for FDA's Office of New Drugs

November 7, 2019

Strengthening the Interpretation of Clinical Trial Data

Jitendra Ganju
Ganju Clinical Trials, LLC
jganju@yahoo.com

Joint work with J Ma, X Yu, K Zhou, Y Lin

Convention

- We prespecify the endpoint and analysis method
- **Formal interpretation relies on ONE method**

  What if we guessed wrong?

Example

Same data, two models, different results

	P-value
Model 1	**0.2274**
Model 2	**0.0004**

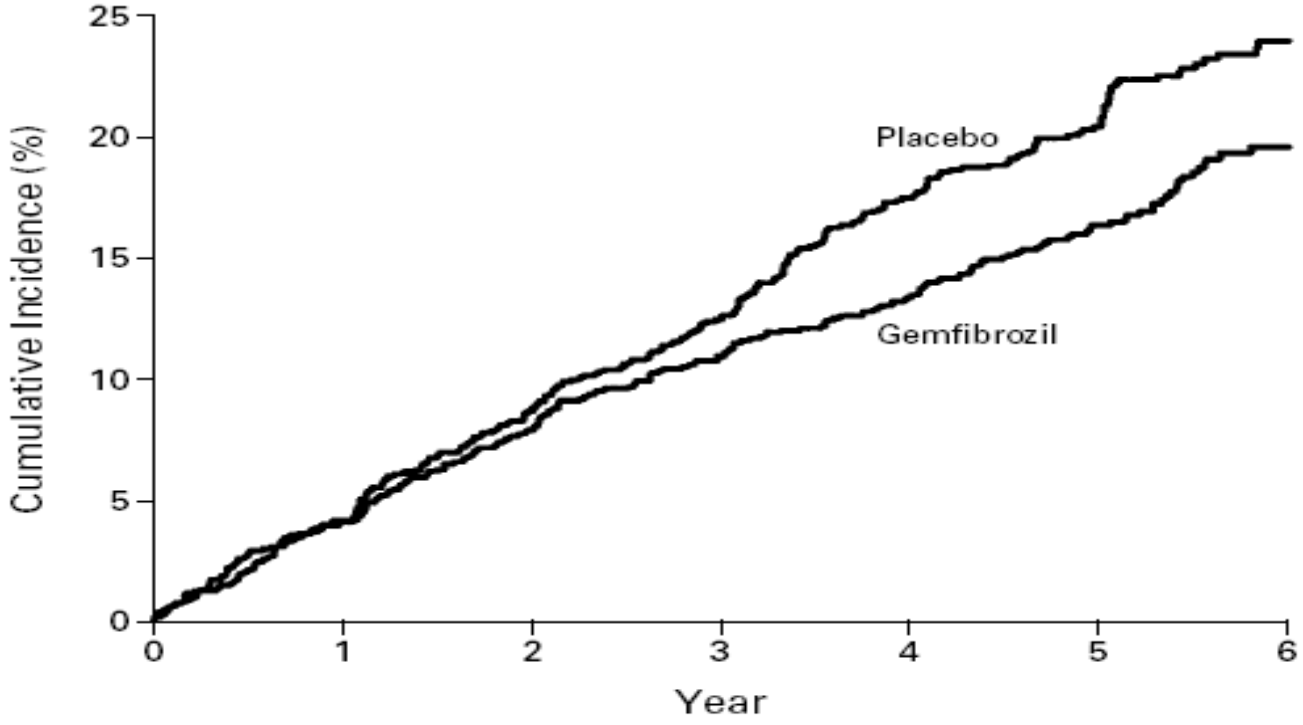
Data from chapter 9 (lots 1 and 2) of *An introduction to the bootstrap*, Efron and Tibshirani (1993)

When single method is risky

- Rare disease, small N
- Complex clinical trials
- Risk in method tied to our experience with endpoint

HbA1c, FEV1,…,6MWT, time to event, recurrent event, new PRO, days hospitalized

Assumption violation



No. at Risk							
Placebo	1267	1200	1118	1040	962	666	466
Gemfibrozil	1264	1201	1128	1067	1005	706	498

Figure 2. Kaplan–Meier Estimates of the Incidence of Death from Coronary Heart Disease and Nonfatal Myocardial Infarction in the Gemfibrozil and Placebo Groups.
The relative risk reduction was 22 percent (P=0.006), as derived from a Cox model.

Rubins et al. NEJM 1995

Proposal

- Prespecify more than one method
- Combine p-values. Control alpha

  Robust, more power, flexible

Robustness

Covariate transformation?

Model	P-value
log(X)	0.2274
No transformation	0.0004
Combined	**0.0040**

Data from chapter 9 (lots 1 and 2) of *An introduction to the bootstrap*, Efron and Tibshirani (1993)

Robustness

Endpoint transformation?

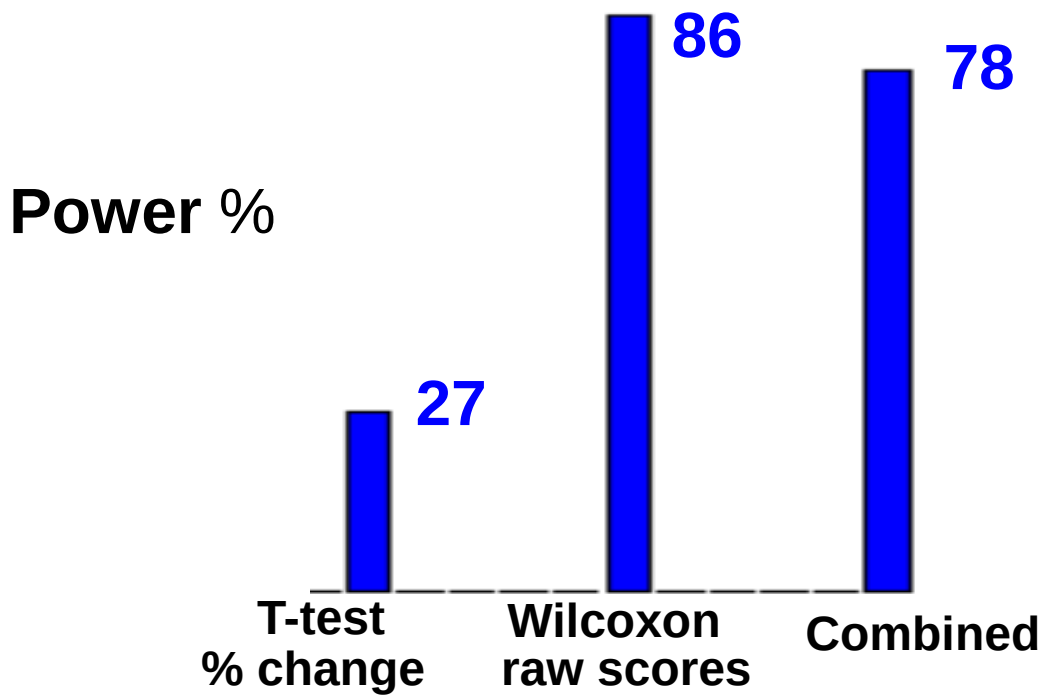
Model	P-value
log(Y)	0.02
No transformation	0.09
Combined	**0.03**

Data from Edwards, Stat Med 1999. Full model fit in each case

Robustness

Different metrics and analysis methods

% change from baseline or raw scores?



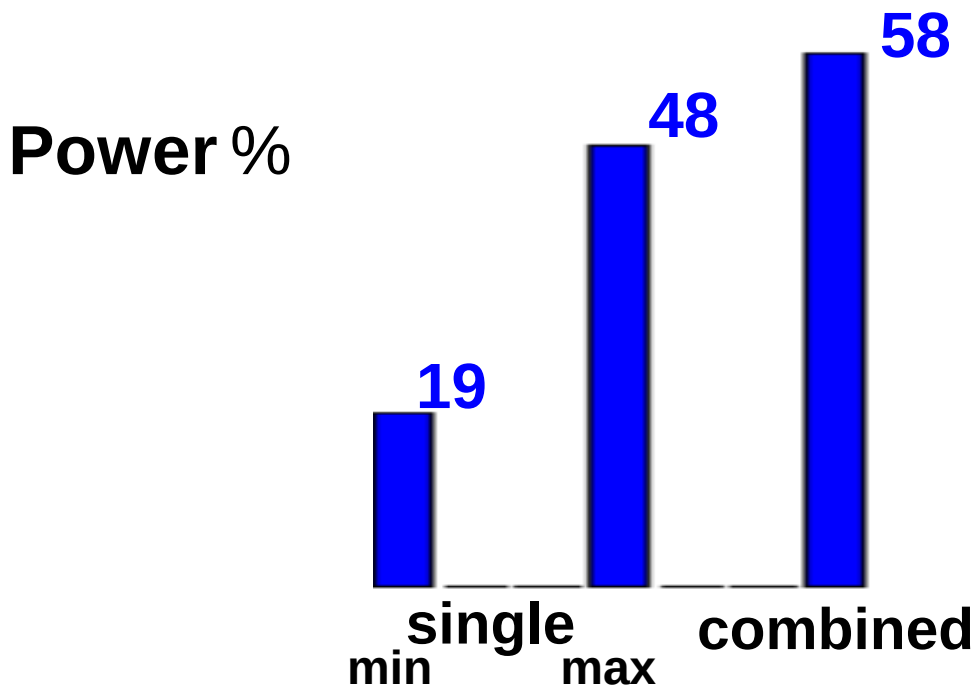
Endpoint is count data

Data from a mixture of Poisson distributions

More Power

Small N, many covariates

N = 20, covariates = 16



Combined method gives more power than any single method

Combined includes 3 methods: one with lowest power, and the other 2 include different subsets of covariates

Versatility

- **Group sequential trials**

  Convention: Same single method at each interim analysis

  Combined methods more flexible

- It’s not just interpretation, trial may stop earlier

Versatility

Different methods at interim and final, and multiple methods at each time

Example

	Convention	New 1	New 2
Interim	LR	wLR	LR, wLR
Final	LR	LR	Cox1, Cox2

As before, combined methods robust

LR = logrank, wLR = weighted logrank, Cox1 and Cox2 are different Cox models

Remarks

- **Limitation:** combining p-values method doesn't give estimate of treatment effect

- To build experience, can start using as complementary method

  Method applies to efficacy or safety endpoints

- Many ways to combine: e.g. min p-value, Fisher's combination
  - Alpha control is via permutations

References

Robust inference

Ganju et al. *Pharm Stat,* 12: 282-290*,* 2013. Correction: *Pharm Stat* 2016

More power

Ganju and Ma. *Stat Methods Med Res*, 26: 64-74, 2014

Group sequential trials

Ganju et al. *Pharm Stat*, 16: 167-173, 2017

Back-ups

Assumption violation

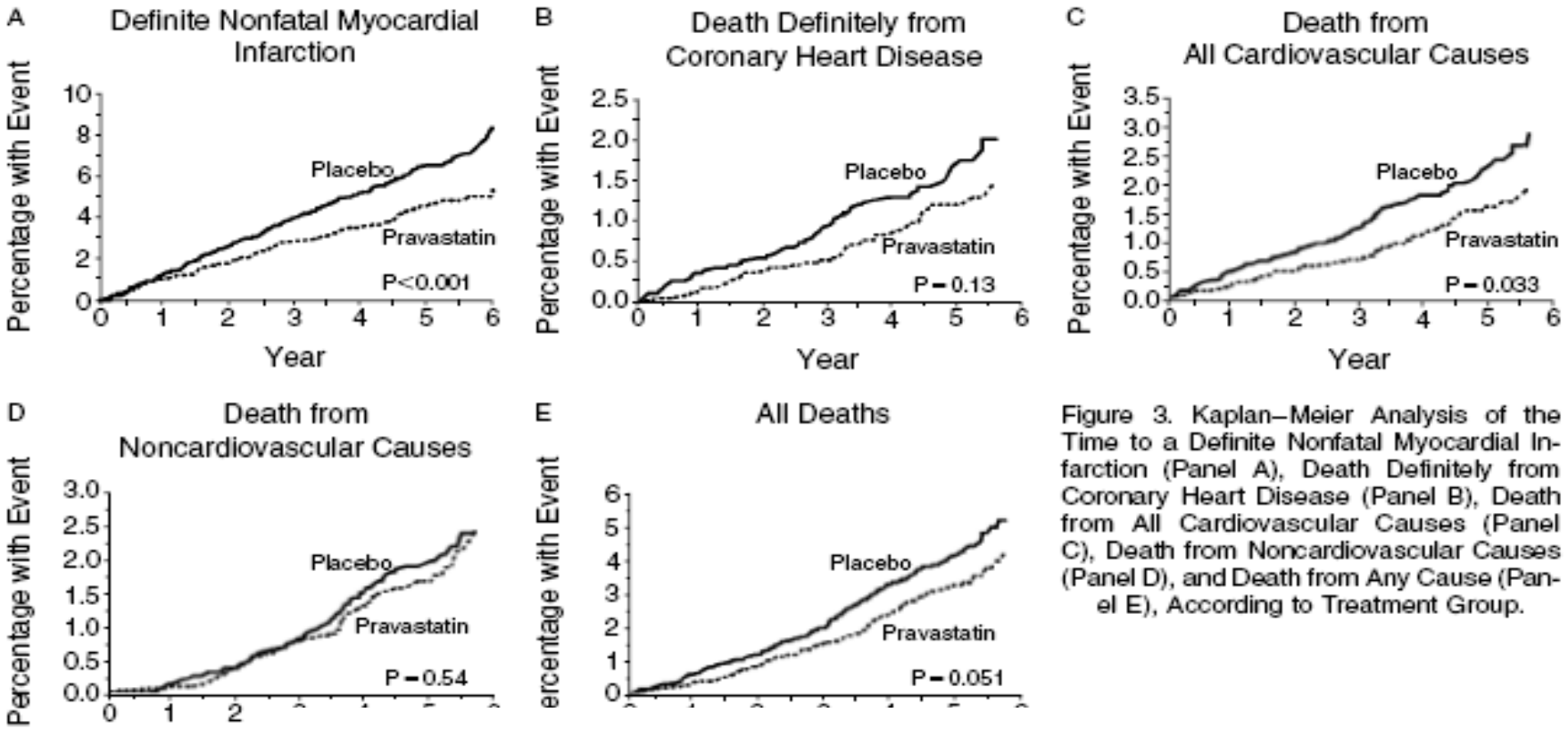
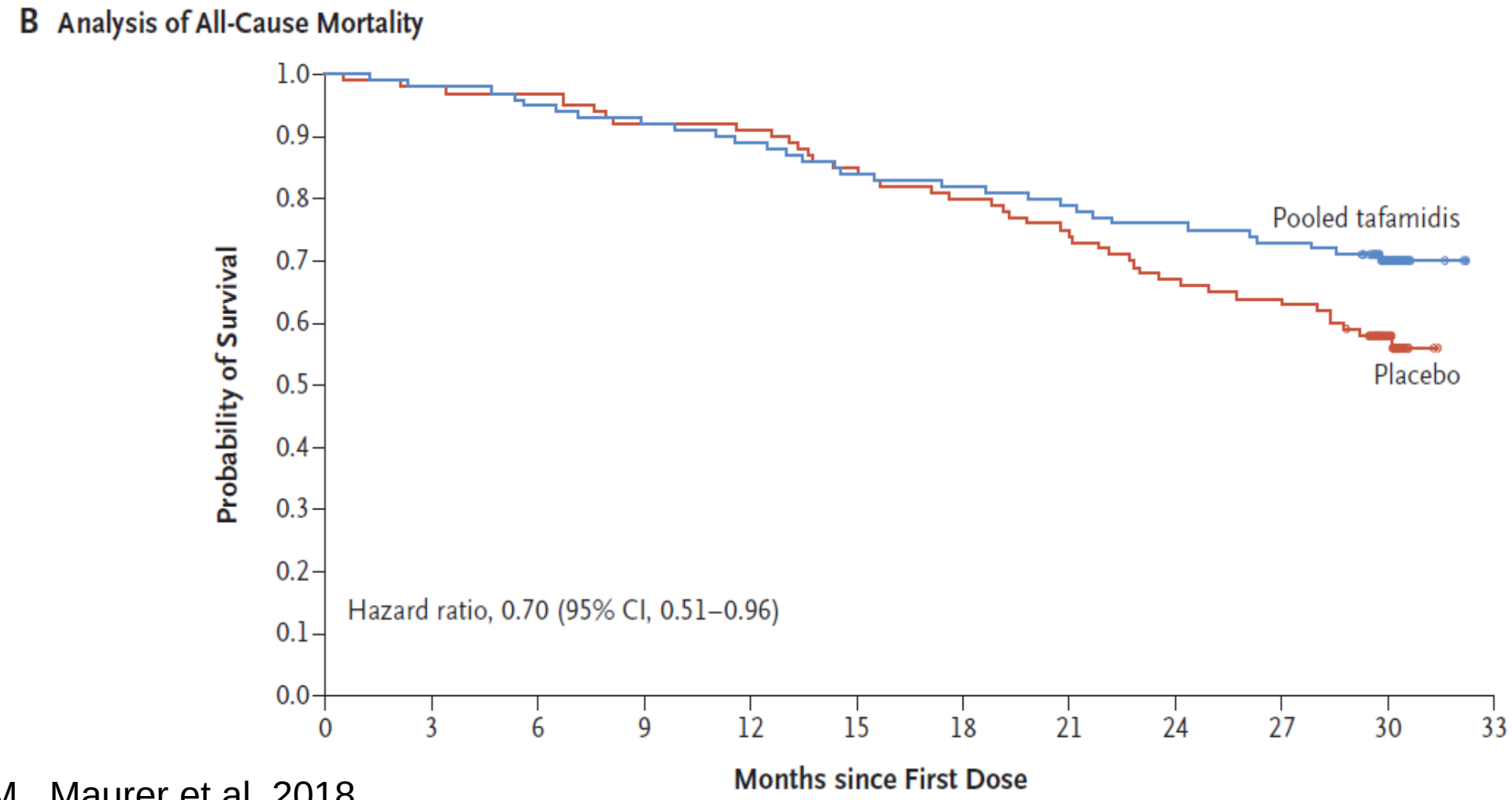


Figure 3. Kaplan–Meier Analysis of the Time to a Definite Nonfatal Myocardial Infarction (Panel A), Death Definitely from Coronary Heart Disease (Panel B), Death from All Cardiovascular Causes (Panel C), Death from Noncardiovascular Causes (Panel D), and Death from Any Cause (Panel E), According to Treatment Group.

PREVENTION OF CORONARY HEART DISEASE WITH PRAVASTATIN IN MEN WITH HYPERCHOLESTEROLEMIA

NEJM , Shepherd et al, 1995

Assumption violation



NEJM , Maurer et al, 2018

Robustness

Endpoint: % change from baseline in Disability Index of Health Assessment Questionnaire in patients with rheumatoid arthritis

Model	P-values
t-test	0.14
Wilcoxon	0.01
Combined	**0.04**

Data from RCT using subset of trial data. N ≈ 60/group