Biostatistics Contributions to Modeling & Simulation in Drug Development

Sue-Jane Wang, Ph.D.

Associate Director for Pharmacogenomics and Adaptive Design
Biostatistics Leader, CDER Biomarker Qualification Program
Office of Biostatistics, Office of Translational Sciences, CDER
U.S. Food and Drug Administration

Presented at Session I: FDA perspectives on applying modeling & simulation in regulatory decision-making: successes and lessons learned, at the "Modeling & Simulation for Medical Products Workshop", Washington DC, September 26, 2013

Acknowledgments

Members of statistical review teams across multiple divisions in Office of Biostatistics contributed their regulatory science research work and their IND/NDA/BLA reviews that have significant impacts on drug labeling as a result of performing modeling and/or simulation

Disclaimer

Except those impacted drug labeling, this presentation reflects the views of the author and those who contributed to regulatory science research, which should not be construed to represent the views or policies of the U.S. Food and Drug Administration

OUTLINE

- Drug development and statistical M&S
- M&S that impacts drug labeling
- M&S that facilitates/enhances regulatory science and statistical reviews
- Predictive M&S: pharmacogenomics and adaptive design planning
- Modeling to address safety risk of drug class
- Concluding remarks

Necessary Experiments in Clinical Development for Regulatory Approval

- Exploratory studies
  - Accumulates preliminary knowledge or root-cause analysis
  - Generates clinical questions from available data (less or more)
  - Not confirmatory analyses
- Confirmatory studies
  - Primary study objectives?
  - Choices of rigorous study designs
  - When effect is demonstrated, further exploratory analyses are often performed, though some may be unplanned

Wang SJ, M&S 2013

What is Statistical M&S?

- Statistical model
  - An approximation to a real phenomena
  - Generally includes a stochastic component
- Simulation
  - Assumed true state of nature
  - Allows step-by-step tracing when the model is executed
- Modeling and Simulation
  - Identifies necessary input data and collect them early
  - Includes appropriate levels of detail for model building
  - Plan for model validation and model verification
  - Plan for statistical analyses, their report and simulation reports

OB M&S At a Glance

- Drug development and statistical M&S
- M&S that impacts drug labeling
- M&S that facilitates/enhances regulatory science and statistical reviews
- Predictive M&S: pharmacogenomics and adaptive design planning
- Concluding remarks

Cardiovascular and Renal

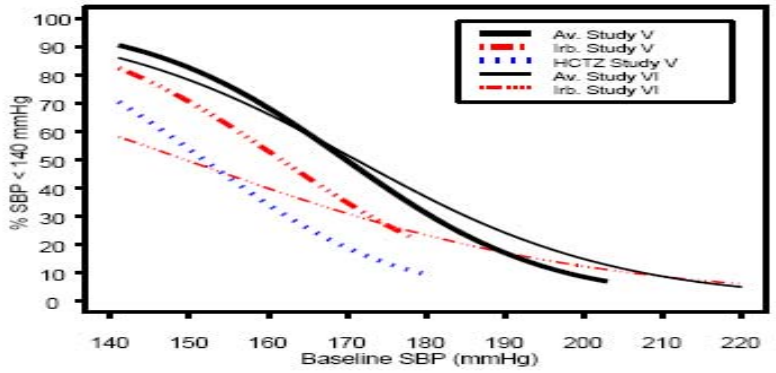


Figure 3a: Probability of Achieving SBP <140 mmHg in Patients from Initial Therapy Studies V (Week 8) and VI (Week 7)*

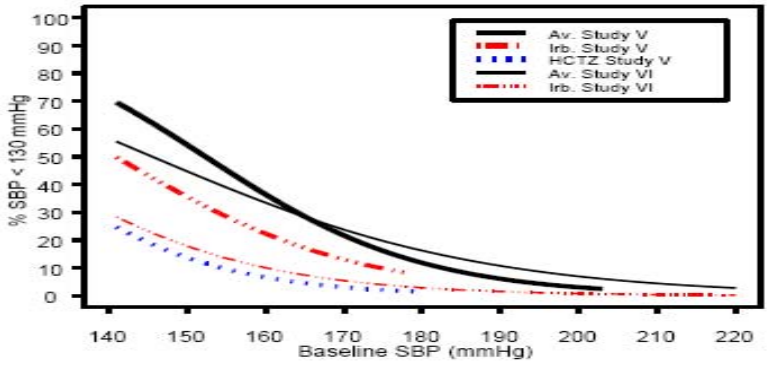


Figure 3b: Probability of Achieving SBP <130 mmHg in Patients from Initial Therapy Studies V (Week 8) and VI (Week 7)*

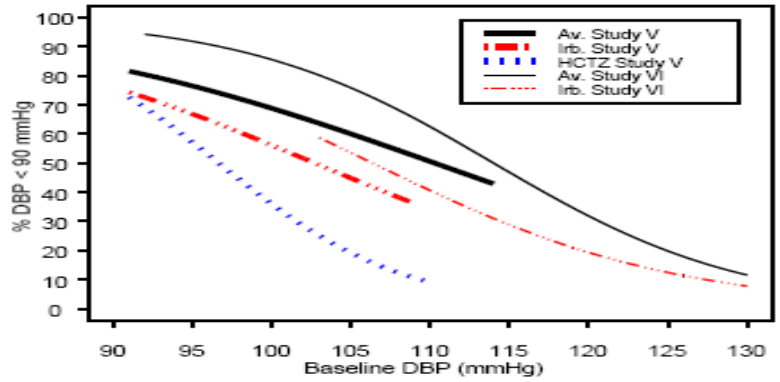


Figure 4a: Probability of Achieving DBP <90 mmHg in Patients from Initial Therapy Studies V (Week 8) and VI (Week 7)*

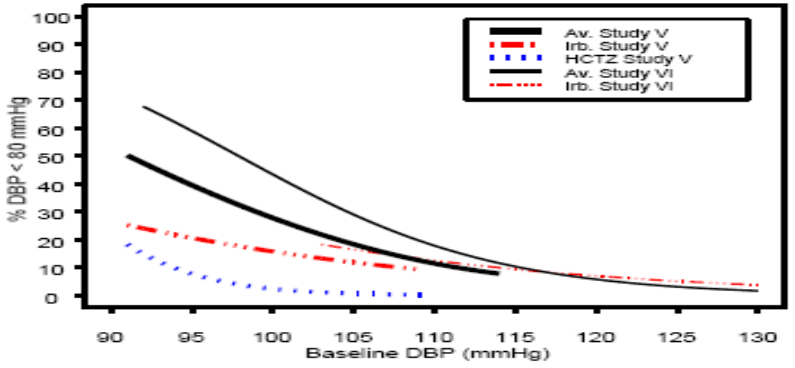


Figure 4b: Probability of Achieving DBP <80 mmHg in Patients from Initial Therapy Studies V (Week 8) and VI (Week 7)*

Wang SJ, M&S 2013

Anti-hypertensive labeling, e.g., Avalide

Cardiovascular and Renal

- Graphics based on proper statistical modeling may help to convey useful information to consumers and prescribers, e.g., probability of achieving normal blood pressure based on baseline BP
  - Importance of model diagnostics
  - Statistical points to consider
  - Choice of cutoffs to avoid misleading graphical presentation
- Profound public health impacts
  - Transition from monotherapy to combination
  - A routine display in anti-hypertensive drug products label

Hung H.M.J., Zhang J, Bai S, Lawrence J Wang SJ, M&S 2013

Model Diagnostics

- Model checking and diagnostics are critical to demonstrate utility of M&S methods, particularly for pivotal (or confirmatory) trials

- What kind of diagnostics are necessary:
  - Goodness of fit test, lack of fit test, ….
  - Model residuals assessment
  - Influence assessment

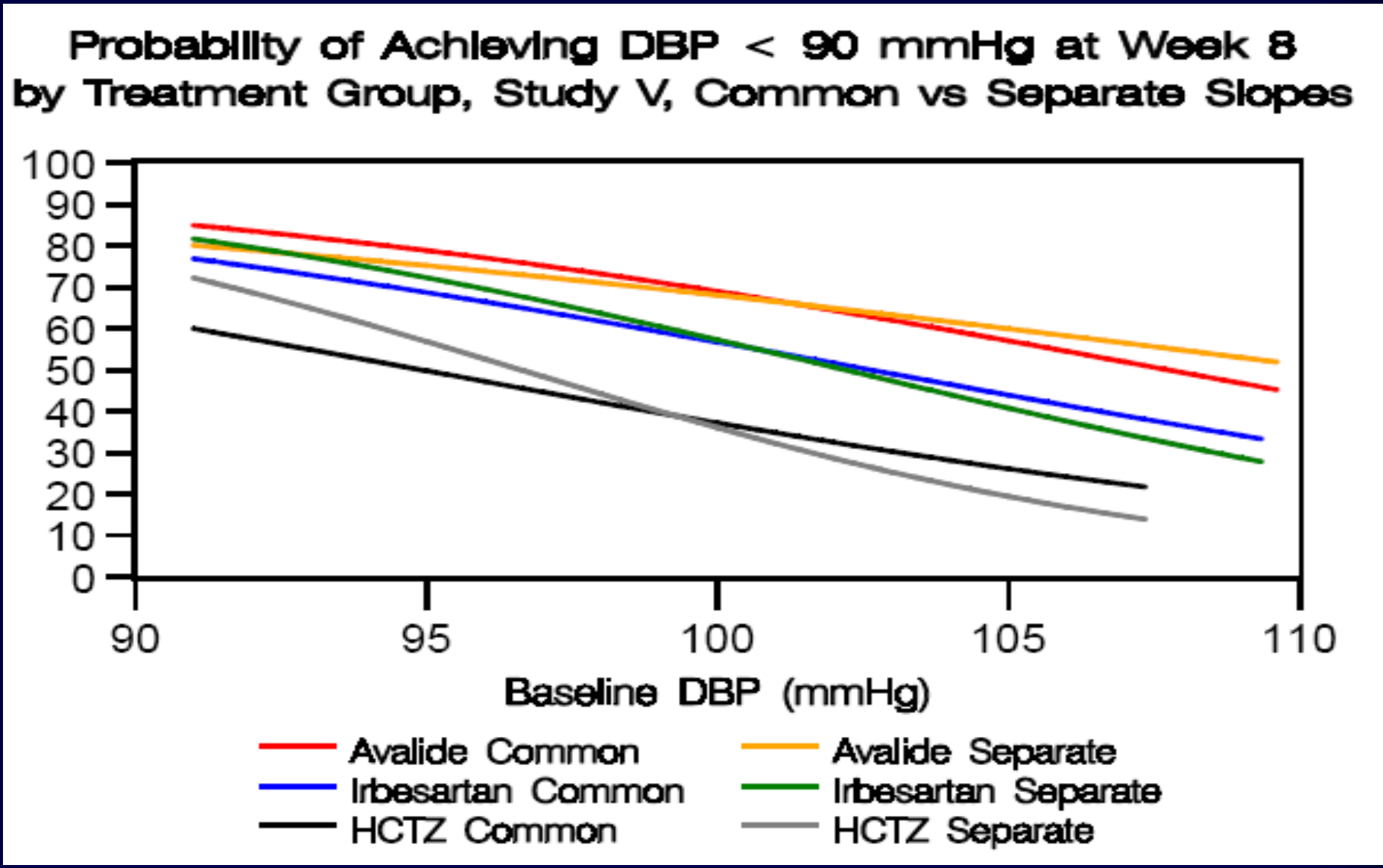
e.g., Hung H.M.J.

Wang SJ, M&S 2013

Table 2.3: Model Diagnostics for Achieving SBP <130 mmHg					
Study/Model	-2 Ln L	Difference	P-value	Hosmer-Lemeshow	P-value
V/Common	507.080	0.334	0.85	5.821	0.67
V/Separate	506.746				
VI/Common	580.509	1.221	0.27	8.340	0.40
VI/Separate	579.288				

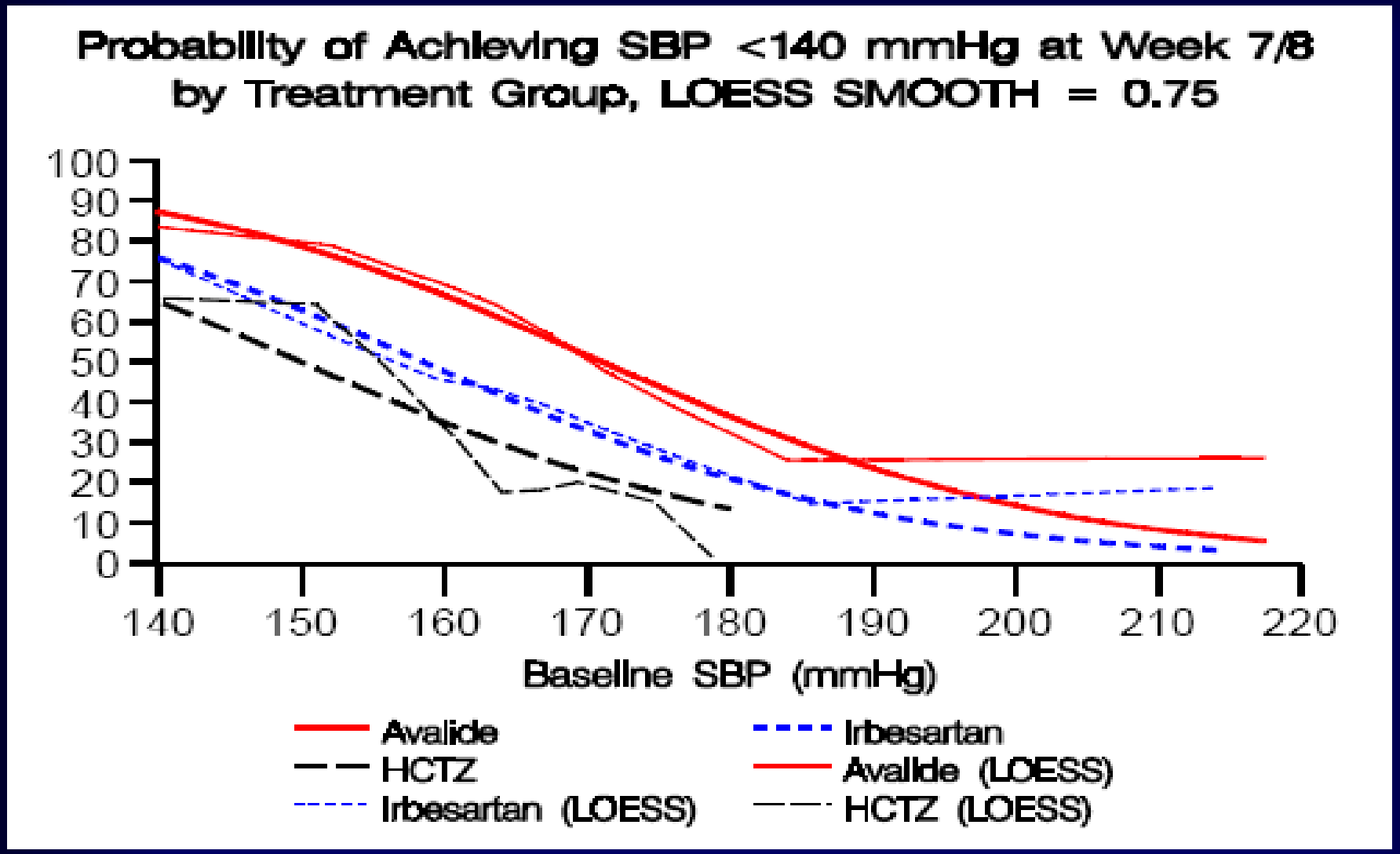
Hung H.M.J., Zhang J

Wang SJ, M&S 2013



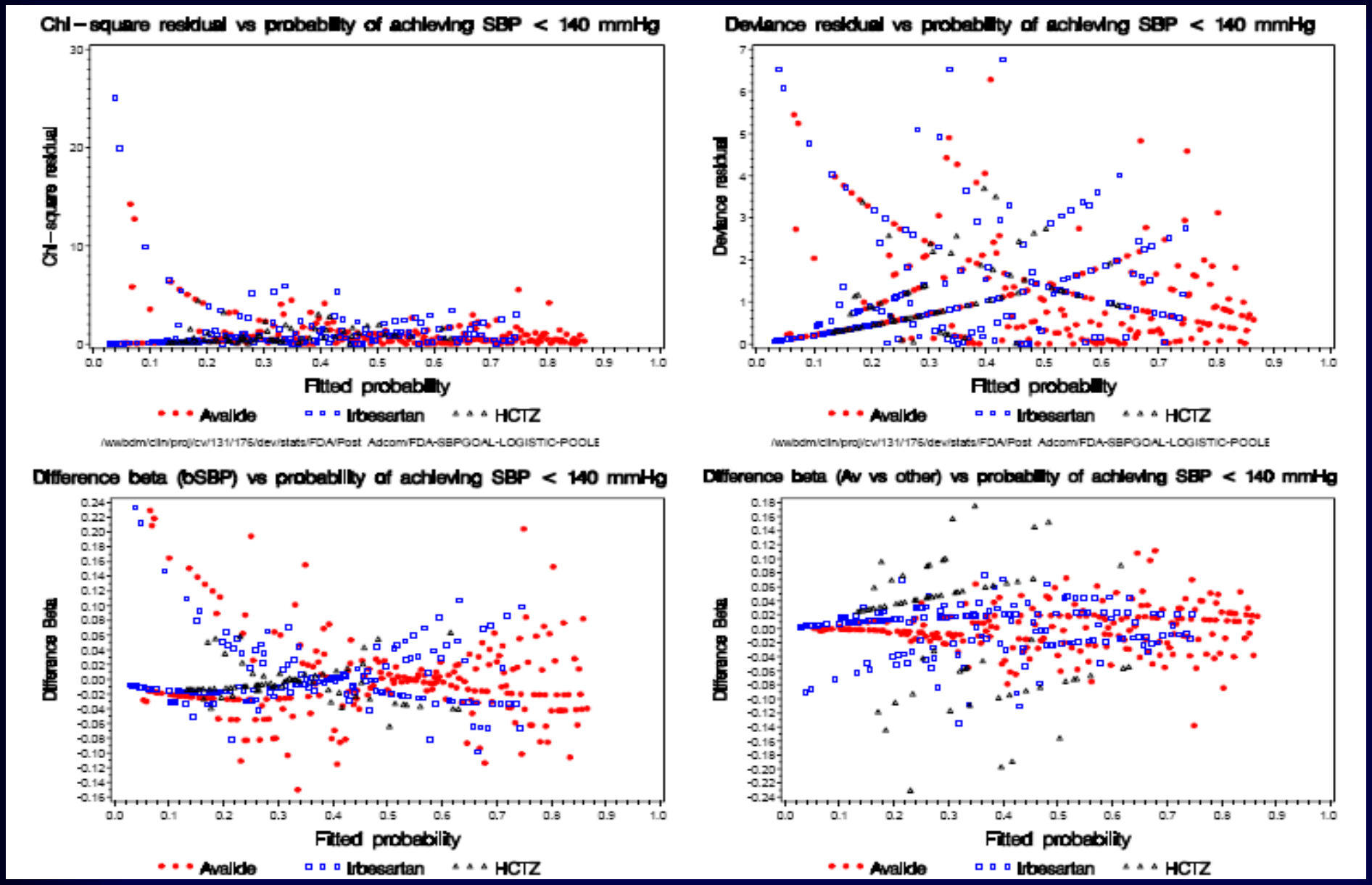
Hung H.M.J., Zhang J

Wang SJ, M&S 2013



Hung H.M.J., Zhang J

Wang SJ, M&S 2013



Hung H.M.J., Zhang J

Wang SJ, M&S 2013

M&S to Illustrate What Might a p-Value of 0.0008 Mean in 2-Arm NI-Trial for Across Trial Inference Relative to Putative Placebo?

Table. Type I error rate associated with comparison of T with virtual P

External Trials			A-C Trial		Empirical Type I error rate		
π_p	π_c	Δ	π_c	π_t	ICI(point e)	ICI (wst limit)	Virtual
K=2, n=5000,5000							
0.20	0.15	0.088	0.15	0.20	0.00006	< 0.00001	0.071
0.20	0.15	0.088	0.13	0.20	< 0.00001	< 0.00001	0.0006
0.20	0.15	0.088	0.17	0.20	0.092	0.020	0.48
K=2, n=1000,9000							
0.20	0.15	0.088	0.15	0.20	0.00059	0.00007	0.089
0.20	0.15	0.088	0.13	0.20	< 0.00001	< 0.00001	0.0016
0.20	0.15	0.088	0.17	0.20	0.11	0.016	0.47
K=10, n=1000 per trial							
0.20	0.15	0.088	0.15	0.20	< 0.00001	< 0.00001	0.028
0.20	0.15	0.088	0.13	0.20	< 0.00001	< 0.00001	< 0.0001
0.20	0.15	0.088	0.17	0.20	0.040	0.0018	0.68

sample size for Active Control trials = 10000 per treatment group

Indirect CI method using 50% rule; based on 100,000 runs; Δ Standard error of the random effect

Wang SJ, M&S 2013

Wang S (presented at 2000 cardio-renal advisory committee meeting), Rashid M

Neurology and Psychiatry

- Impacts of dropout models, correlation structure, analysis approaches with simulated incomplete data versus real data on treatment effect assessment
- Discerning disease modifying drug effects in early Parkinson's disease (OB and OCP collaborative research)
- Modeling treatment effect in MRCT and design issues (OB&ODE-II)
- Enrichment design in placebo non-responders – simulation to compare different designs and different missing data imputations
- Comparing novel (e.g., sliding dichotomy) with conventional (e.g., proportional odds) models on operating characteristics for confirmatory evidence, e.g., patients with traumatic brain injury
- Assessing accuracy/appropriateness of simulations done by sponsor
- Exploring population heterogeneity with heavy dropouts

Wang SJ, M&S 2013

Siddiqui O, Hung HMJ, O'Neill RT, Chen YF, Gobburu, VSJ, Wang SJ, Laughren T, Ni AK, Yang Y, Kong F

Assessing Design and Analysis Approaches Clinical, Bioequivalence, Carcinogenicity, Safety

- Simulation studies to compare single-stage vs two-stage design for bioequivalence crossover studies concluded better performance using a single-stage design
- Rank analysis, estimation and associated inference on
  - General mixed model w/ covariates in multicenter trials
  - Unbalanced repeated measures model with incomplete repeated measures data
- Simulation studies to compare and propose alternative design for long term rodent carcinogenicity studies
- Bayesian analysis to drug safety for the support of no safety concern versus having safety concerns

Wang SJ, M&S 2013

Rashid M, Li H, Grosser S, Jackson MT, McEvoy B, Tiwari R

Oncology and Hematology

- To understand operating characteristics including bias, uncertainty and impact on false positive and false negative errors
  - Frequency of progression assessment by tumor types defined by median PFS
  - Among (i) methods of censoring or counting the event when disease progression occurs after two or more missed assessments and (ii) assessment interval relative to median TTP
  - Asymmetry in progression assessments between T and C (or P) due to (i) delay by arm, (ii) variability in delay ➔ right censoring more bias than interval censoring

Wang SJ, M&S 2013

Zhang L, Ko CW, Tang S, Sridhara R, Jiang X, He K, Chattopadhyay S, Holland C

Time-to-event Data, Group Sequential Trials, Accelerated Approval

- Assessing the degree of wrongly concluding an ineffective treatment effective by varying censoring rates, median PFS, and censoring patterns due to differential toxicity
- Illustrate via simulation studies the phenomenon of over-estimation of effect size in group sequential trials
- Simulation studies to explore relationship between PFS & OS when crossover is an option in oncology applications: rate of crossover & survival post progression lengths can confound OS effect
- Explore ad-hoc rules to estimate probability of incorrect accelerated approval decision by varying hazard ratio on TTP or OS with pre-specified interim time in a group sequential setting via simulation

Wang SJ, M&S 2013

Yuan V, He K, Jiang J, Sridhara R, Zhang JJ, Tang S, Yang P, Chen G, Chi G, Norton J

Predictive Modeling and Simulation, Pharmacogenomics and Adaptive Design

- Predictive modeling and simulation from high dimensional data for classifier development and explore impacts due to misclassification
- Modeling using NDA/BLA data to identify important factors including Pharmacogenomics biomarkers associated heterogeneity
  - Superiority studies versus non-inferiority studies
- Nested adaptive enrichment
- Modeling and simulation to explore operating characteristics with an adaptive design
  - Exploratory studies
  - Confirmatory studies

Wang SJ, M&S 2013

Wang S collaborative regulatory research science publications

Modeling to Help Quantify the Increased Risk of Suicidal Behavior or Ideation for Patients & Prescribers

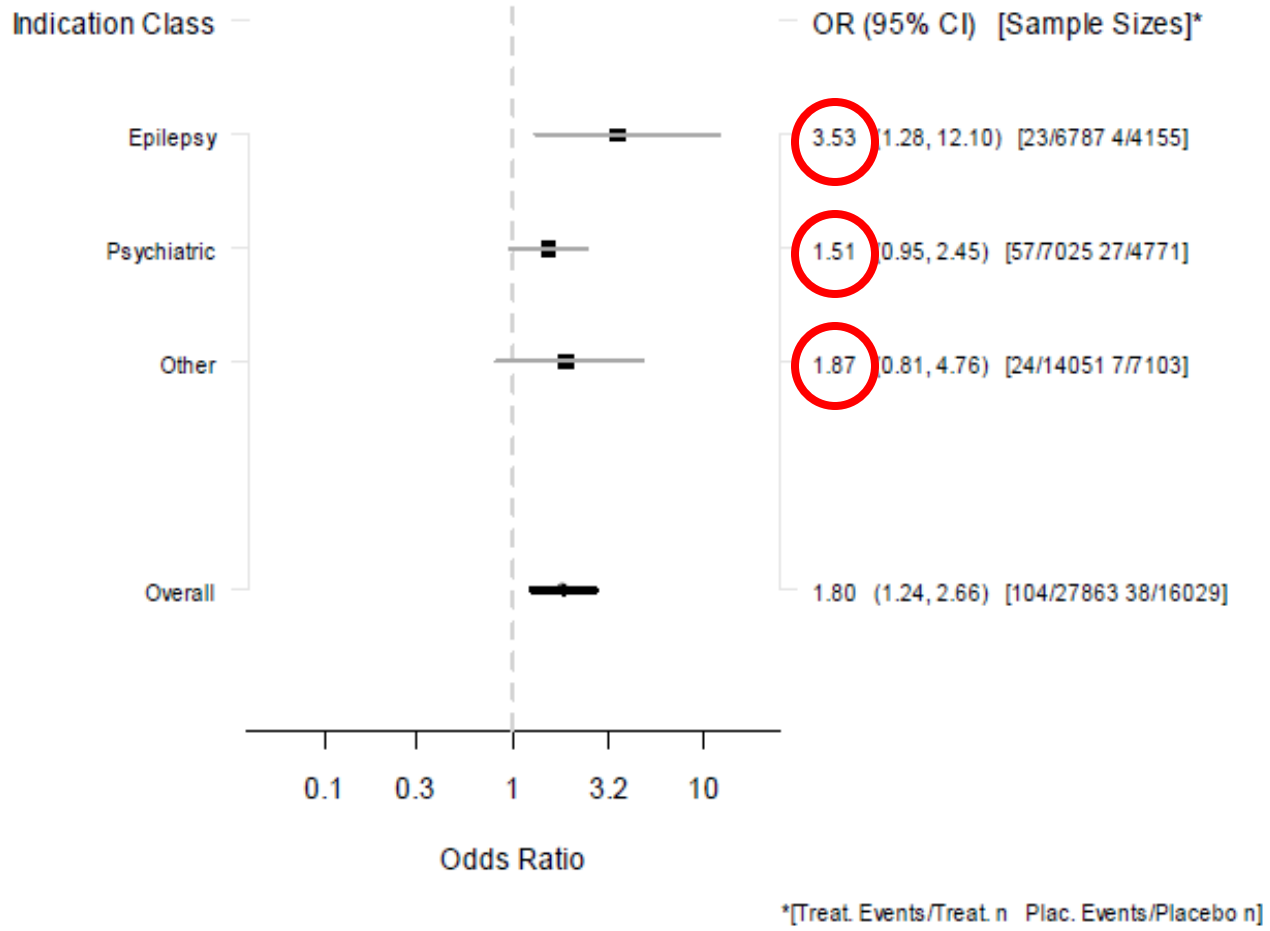


Figure 8: Suicidal Behavior or Ideation Odds Ratio Estimates by Indication Group, Placebo-Controlled Trials.

Wang SJ, M&S 2013

Levenson M, Rochester CG – Biostatistics review and evaluation

Section 5.4: AEDs increase the risk of suicidal thoughts or behavior in patients taking these drugs for any indication

A generalized linear mixed model (GLMM) (McCulloch and Searle 2001) was used to estimate the overall odds ratio in the presence of trial heterogeneity of the odds ratio. The model used the binomial error distribution and logit link function. The model included fixed effects for the trial and treatment effects and a random effect on the trial-level for the treatment-trial interaction. The estimate and the 95% confidence interval of the treatment effect were qualitatively compared to those from the primary method to examine the effect of trial heterogeneity. The confidence interval of the variance component of the random effect was also examined to evaluate trial heterogeneity.

Suicidal Behavior and Ideation and Antiepileptic Drugs

Update 5/5/2009:

AED class label changes

Manufacturers of antiepileptic drugs (AEDs) or anticonvulsant drugs will update product labeling to include a warning about an increased risk of suicidal thoughts or actions and will develop a Medication Guide to help patients understand this risk. These changes affect all approved AEDs except those indicated only for short-term use.

Drugs with updated labels

The approved AEDs affected by these safety label changes are Carbatrol, Celontin, Depakene, Depakote ER, Depakote sprinkles, Depakote tablets, Dilantin, Equetro, Felbatol, Gabitril, Keppra, Keppra XR, Klonopin, Lamictal, Lyrica, Mysoline, Neurontin, Peganone, Stavzor, Tegretol, Tegretol XR, Topamax, Tranxene, Tridione, Trileptal, Zarontin, Zonegran, and generics. FDA approved updated labeling for these drugs on April 23, 2009.

Wang SJ, M&S 2013

Levenson M, Rochester CG, http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm100190.htm

M&S in Exploratory Trials

- Informative learning to gain preliminary observation for correct go/no-go decision is critical – adaptive or fixed
- To improve the probability of correct selection based on early phase exploration relies upon
  - Ability to make go decision and not making type II error
- Patient population "starts narrow and ends wide" versus "starts wide but ends narrow"
  - Intended population well-understood vs potential biomarker
- # of dose groups starts a few versus more
  - If adaptive design: exploratory or confirmatory?
  - Use of M&S depends on simulation study design, objectives and level of details including adaptive design

M&S in Confirmatory Studies

- Sufficiently powered
  - M&S used to evaluate Best Statistical Testing Procedure with multiple arms and where dose-response model may be assumed
  - A MCP procedure that is robust to DR model misspecification
- Statistical modeling
  - Modeling used to explain relationship between baseline characteristics, e.g., prognostic or predictive baseline factors, and clinical outcome mostly of primary interests
  - Model diagnostics needed before interpretation of results
- Caution: multiple phase 3 trials
  - Model chasing + improper accounting for missing data and missing mechanism in confirmatory trial – inability to conclude

Wang SJ, M&S 2013

M&S pursued by, e.g., mostly Hung, Wang, and FDA/CDER/OTS/OB biostatistics review scientists

Concluding Remarks

- M&S has great values to gain useful insights from exploratory trials
- For descriptive purposes, modeling requires thorough model diagnostics according to the well-founded statistical principles
- M&S for confirmatory trials is usually performed to evaluate the properties of trial design and statistical analysis strategy
- For testing clinical hypotheses in confirmatory trials, primary statistical tests can be challenging if relying on specific models and assumptions - need to be robust to model assumptions for general consideration
- We have illustrated many more utilities of M&S for drug development through statistical M&S; their impacts on drug labeling continue

BACKUP Slides

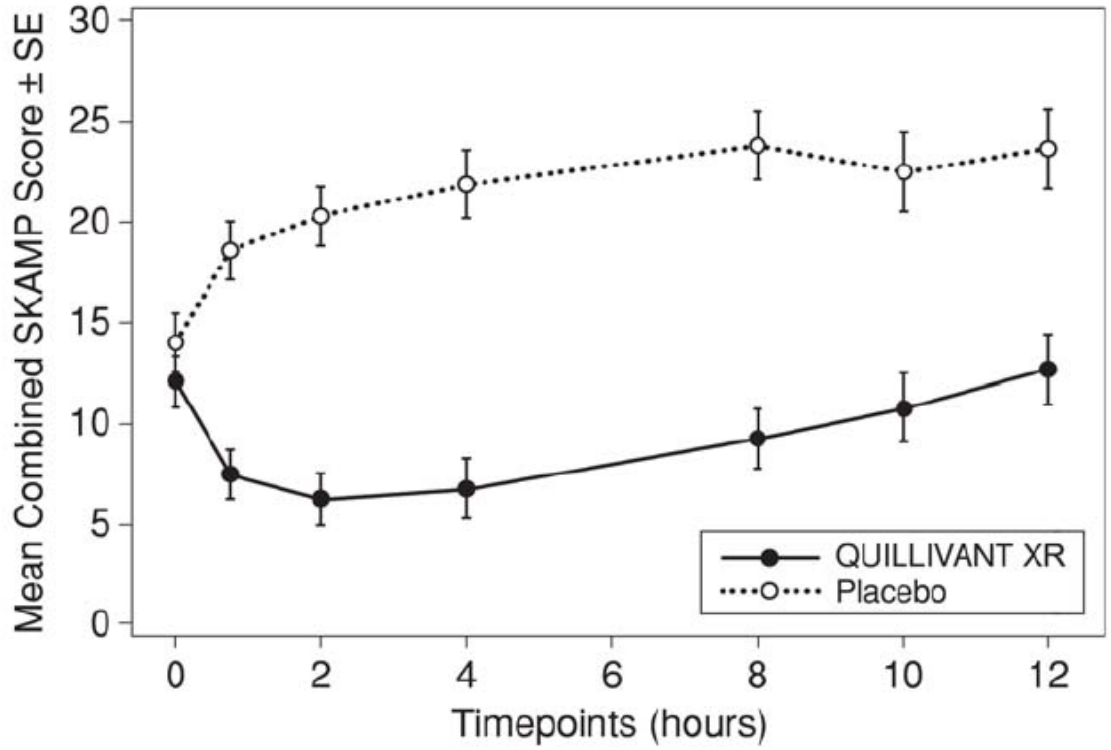
Psychiatry

- A two-period crossover trial design
- Interpretation of study results depends on if there is concern of carryover effect or sequence effect
- To avoid misleading representation in the presence of obvious carry over effect
- A recommendation on how treatment effect over time should be presented

Quillivant XR label for treatment of ADHD; Chen YF, Yang PL, Hung HMJ

Treatment Effect as Time Course Plot in Crossover Trial When Period is Confounded with Differential Carryover or Sequence Effect

Figure 3. Absolute SKAMP-Combined Score after treatment with QUILLIVANT XR or Placebo during Period 1.



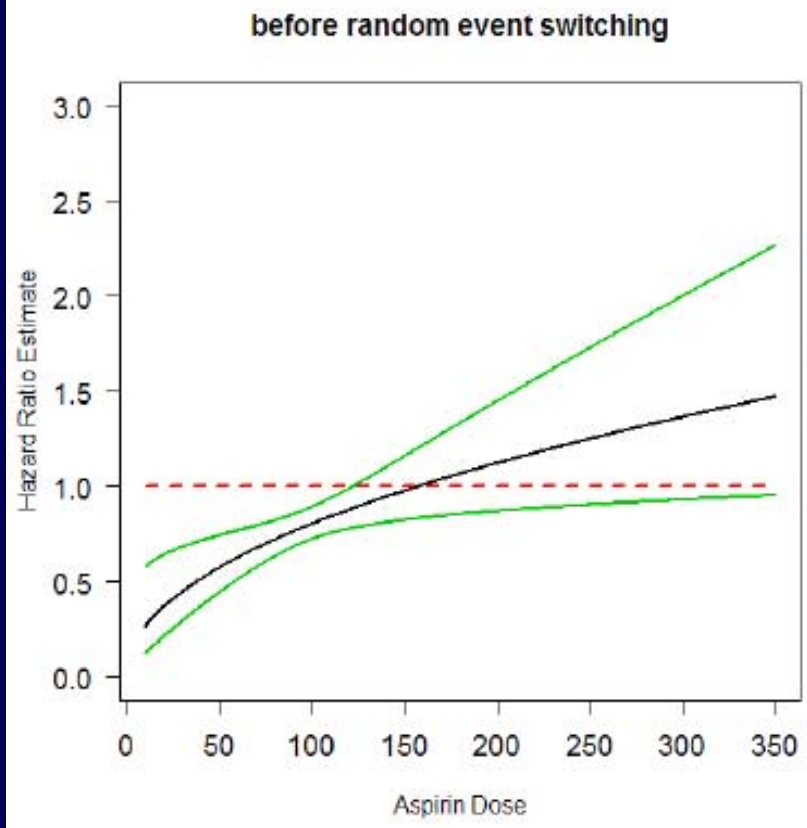
Quillivant XR label for treatment of ADHD; Chen YF, Yang PL, Hung HMJ

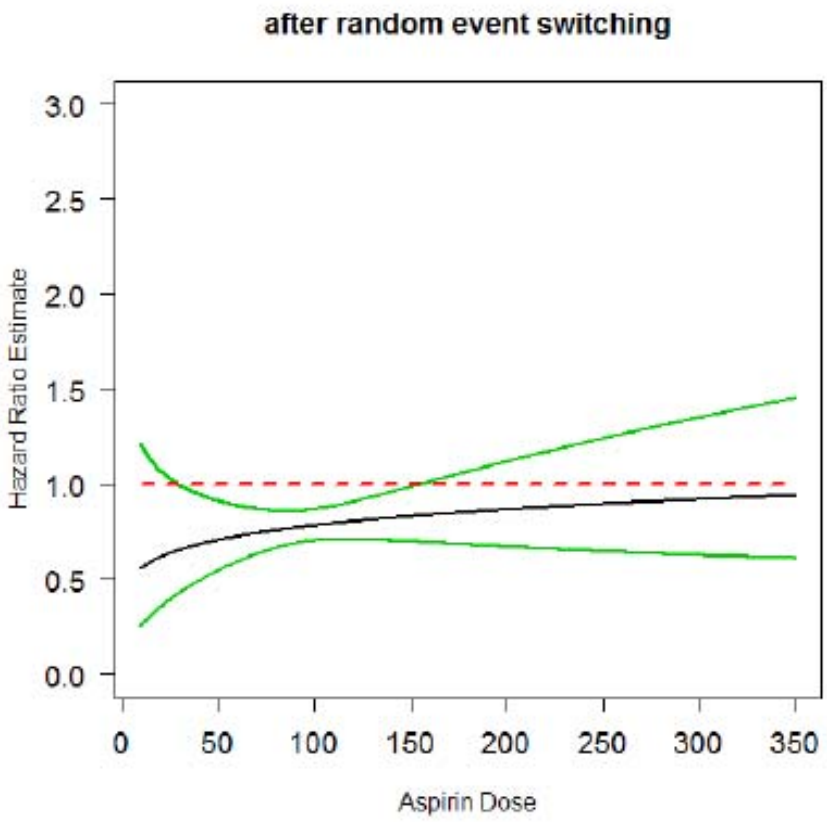
Wang SJ, M&S 2013

U.S. Food and Drug Administration
Protecting and Promoting Public Health

www.fda.gov

The Cox Proportional Hazards Model is Sensitive to the Few Events in High ASA Subjects from Non-US





Wang SJ, M&S 2013

Statistical modeling using PLATO trial data by Zhang J, Hung HMJ, presented at Cardio-Renal AC meeting, 07-28-2010

Conclusions

- Although play of chance can never be ruled out as a possible explanation, the probability of play of chance seems low.
- Concurrent use of aspirin is shown to be the single best factor for explaining the geographic disparity; however, the robustness of the explanatory model is in question.
- The reviewer was not able to find a definitive explanation by looking at other available covariates.

Wang SJ, M&S 2013

Statistical modeling using PLATO trial data by Zhang J, Hung HMJ, presented at Cardio-Renal AC meeting, 07-28-2010