

How Liberal Can Learn-Free Type I Error Rate be In Confirmatory Adaptive Trials

Sue-Jane Wang, Ph.D.

Office of Biostatistics, Office of Translational Sciences
Center for Drug Evaluation and Research, U.S. FDA

Presented at “Adaptive Design Working Group – KOL Series”
May 21, 2010

The views presented are those of the authors and not necessarily those of the US FDA

Collaborators

H.M. James Hung
Robert O'Neill

Acknowledgments

The research was supported by FDA/CDER
Regulatory Science Research Awards

Outline

- Definition
- Perceived Benefits
- Exploratory or Confirmatory
- Learn and Confirm
- Learning-Free Type I Error Rate
- Simulation studies
- Concluding Remarks



Adaptive Design

- **Prospectively planned opportunity** for modification of one or more specified aspects of the **study design and hypotheses** based on analysis of data (usually interim data) **from subjects in the study**
- Analysis of the accumulating study data are performed at **prospectively planned time-points** within the study
- **Analyses** can be performed in a **fully blinded** manner or in an **unblinded** manner, and **can occur with or without formal statistical hypothesis testing**

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf>

Perceived Benefits

The efficiency of a drug development by

– Potentially saving the total study sample size

- Increase the sample size within a trial only when necessary
- Dropping an arm that is likely to be ineffective or intolerable/unsafe
- Principles and rationales of AD in various stages

– Potentially reducing the total investigation time

- Shorten the white space between Ph II (stage 1) & Ph III (stage 2) by combining the planning of both phases and move it to before Ph II implementation: Do we lose critical think time?
- Require upfront resources, logistics, and planning time including performing simulation studies prior to trial initiation

The potential advantage of Adaptive Design is its flexibility with adaptation rules pre-specified.

Where adaptation tries to deal with better learning and formalize the learning, it should not be confused with the necessity to prospectively plan an (adaptive) design clinical trial that is adequate and well-controlled

From Exploratory Phase to Confirmatory Phase: Cannot Skip Learning Step

- Principles of exploratory studies are needed
 - Patient population, subgroups and/or doses
 - Several intermediate efficacy endpoints
 - Limited available data on clinical endpoint(s) to learn drug associated dropout pattern over time
 - Some idea on distribution of clinical endpoints, explore a sensitive definition, its effect size, etc.
 - Dose-response study vs. identification of promising dose-regimen(s), may rely on 'surrogate' endpoint
- Use of point estimate of effect size from phase II to plan ph III can be valuable in prediction of useful doses, but, may be too optimistic with usual α -level and $1-\beta$ level for phase III planning (Wang, Hung, O'Neill; Pharmaceutical Statistics 2006)

When desired to “learn” and “confirm”

Phase II vs. Phase III

Stage 1 vs. Stage 2

Early Aspect vs. Later Aspect

Part 1 vs. Part 2

The purpose is to combine early stage 1 learnt data and stage 2 independent data such that it is a one-trial scenario



Adequate & Well-Controlled (A&WC) 21CFR314.126

- Not exploratory adaptive design clinical trial
- In addition to experimentwise type I error rate control
- Should possess the following characteristics
 - ◆ clear statement of the objectives, proposed and actual methods of analysis in protocol, SAP, and reports
 - ◆ design that permits a valid comparative evidence of T-effect
 - ◆ methods of adequate assurance of patient selection
 - ◆ patient assignments that minimize bias, group comparability
 - ◆ minimize bias on all parties: pts, investigator, data analyst
 - ◆ endpoints well-defined that address clinical primary hypo.
 - ◆ analysis results - interpretability of the effects of drug

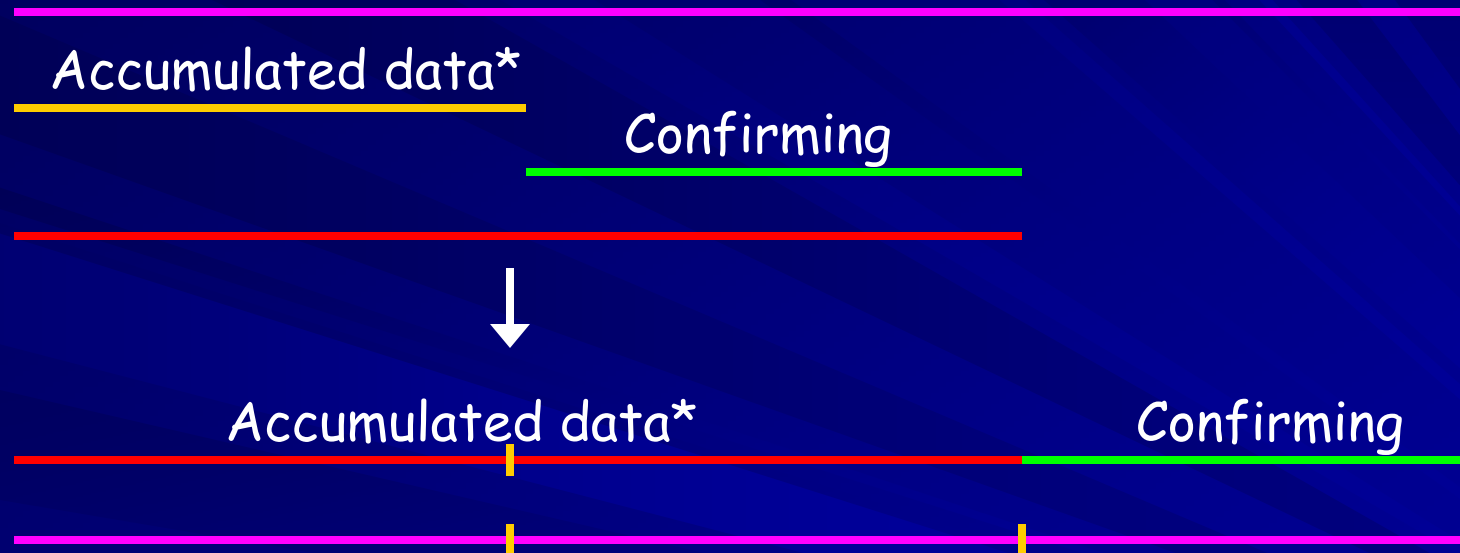
For an A&WC Trial

One Study-wise Type I Error Rate

Group sequential design (GSD) is a special case of adaptive design

- Each endpoint has its own α -spending function
- The α -spending function deals with repeated significance testing properly
- GSD provides flexibility to terminate the study early for (i) efficacy, (ii) futility
- Maximum statistical information (sample size or event number in survival endpoint) is fixed
- Limited trial extension based on blinded assessment is permitted

Learning/Confirming in GSD with a single endpoint: Inference based



*The analyses based on the accumulated data (whether or not calling it learning) are prospectively planned that are inference based i.e., well-defined α -spending poses α -penalty on making the repeated use of the accumulated data (in the sense of learning) for confirming

Dealing with Learning

(i) Conventional GSD that allows minimum changes (n, duration)

Analysis using accumulated data via α -spending making it inferential

(ii) Combine Learning with Confirming (change global H_0 hypothesis)

Flexibility

Learning-Free Type I Error Rate

The type I error rate associated with only the selected final hypotheses in stage 2 without considering necessary adjustment of type I error rate associated with use of the learning data, equivalently, without subjected to the needed multiplicity adjustment for a family-wise type I error rate control

In the simplest case:

When learning does not
involve extending the total
sample size

When learning data is used to select a treatment arm mid-trial

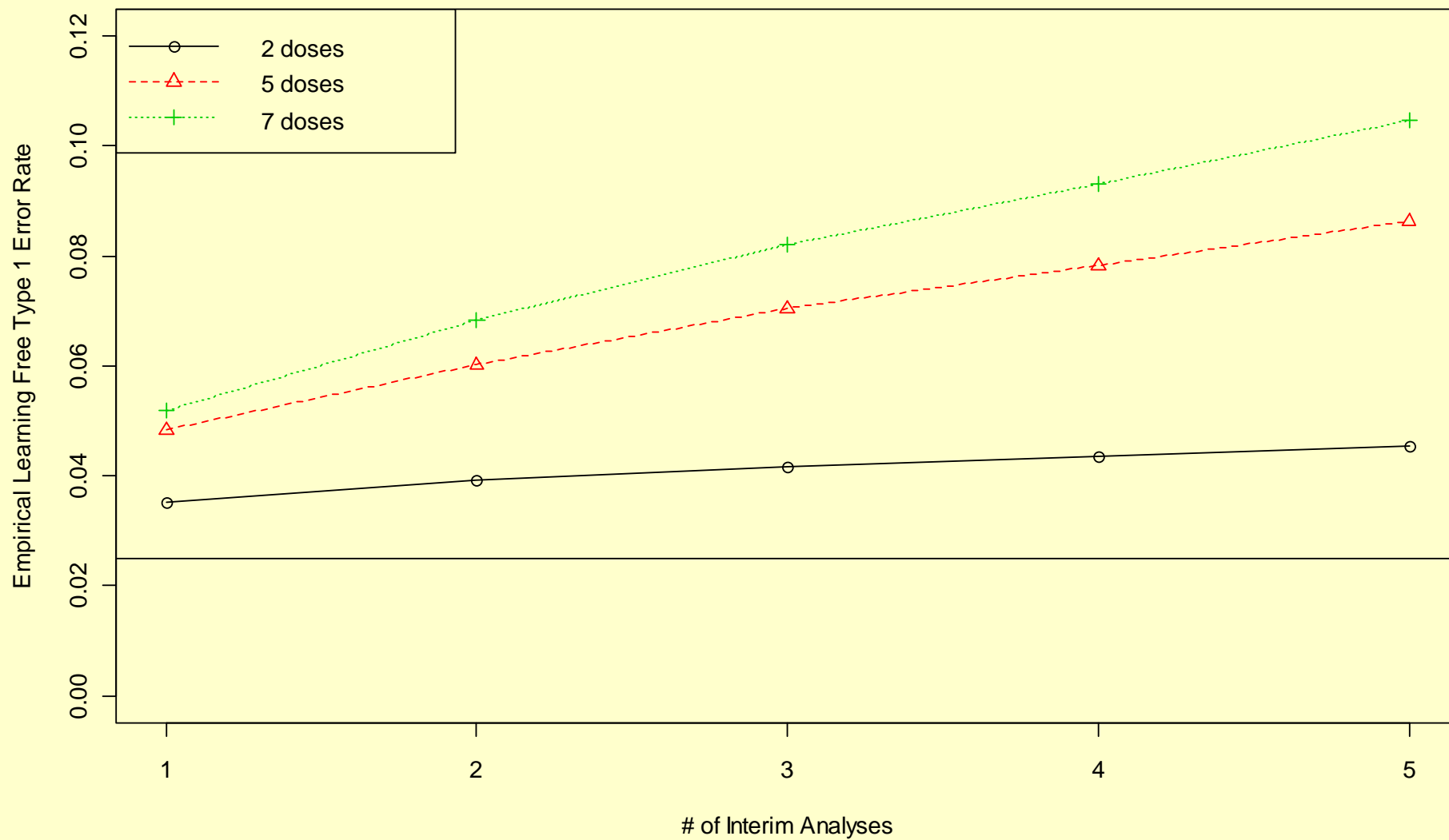
Select Among Dose Regimens Only

For example:

If standardized difference is at least "D" SD for all $j = 1, \dots, J$, but max-group

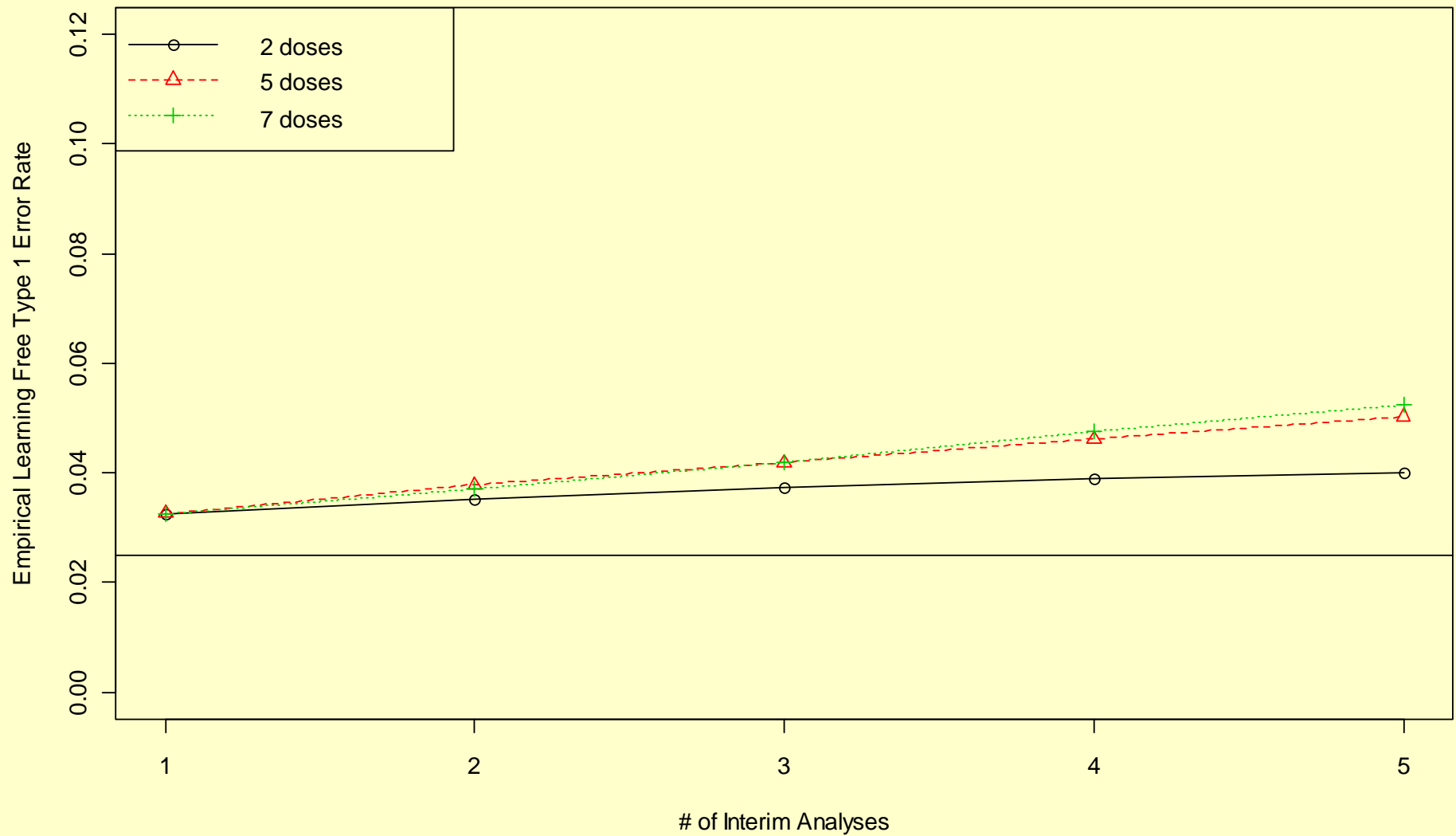
- The only selected dose is compared to placebo in the final analysis at α level, and the remaining dose regimens are dropped
- For other scenarios, the final analysis may be tested at α / k significance level with $k \leq J$

Figure 1. Select Among Treatment Dose Groups Based on Clinical Endpoint, Cutoff $D = 0$



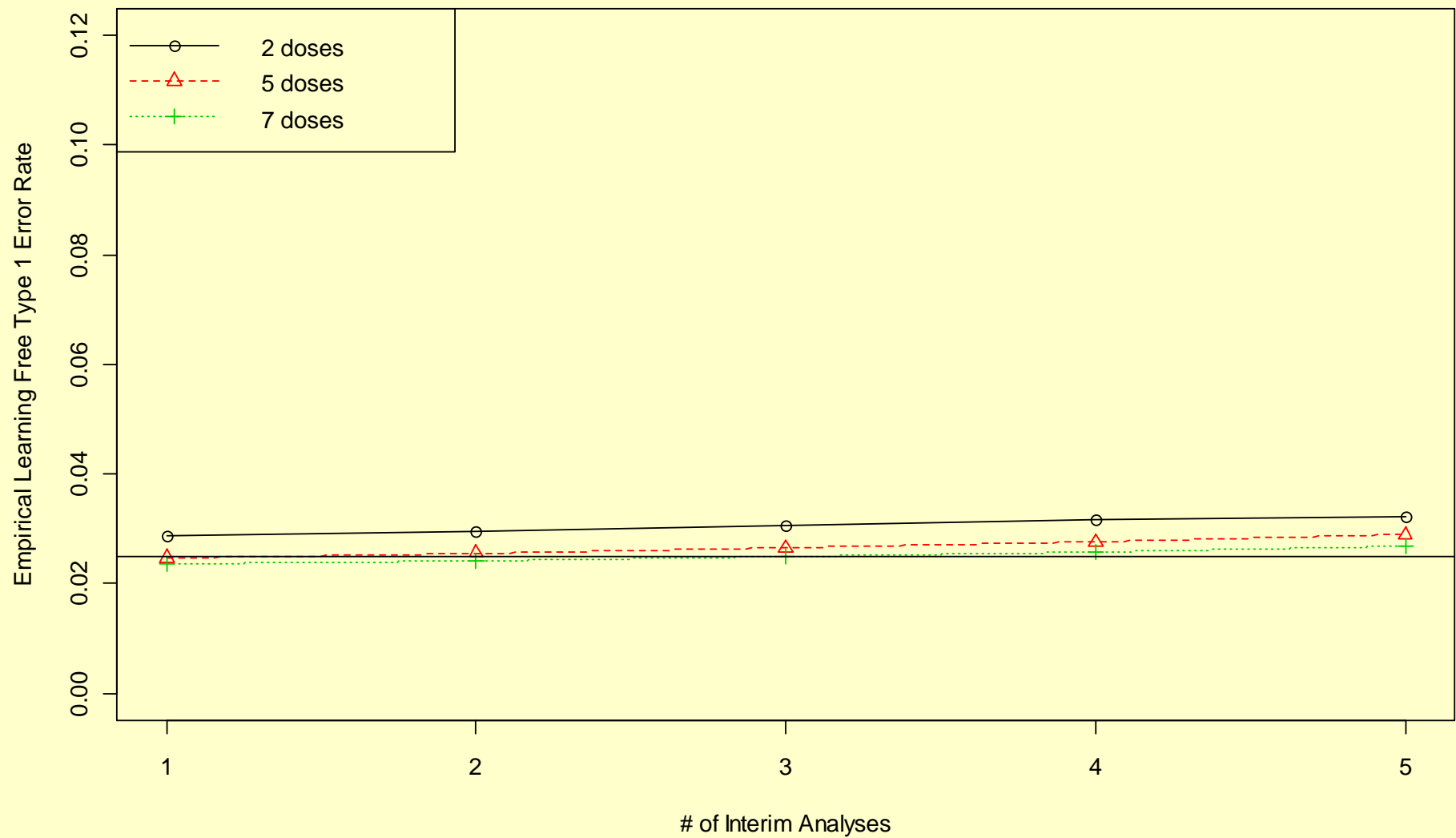
Wang, Hung, O'Neill (2010)

Figure 2. Select Among Treatment Dose Groups Based on Clinical Endpoint, Cutoff $D = 1$



Wang, Hung, O'Neill (2010)

Figure 3. Select Among Treatment Dose Groups Based on Clinical Endpoint, Cutoff D = 2



Wang, Hung, O'Neill (2010)

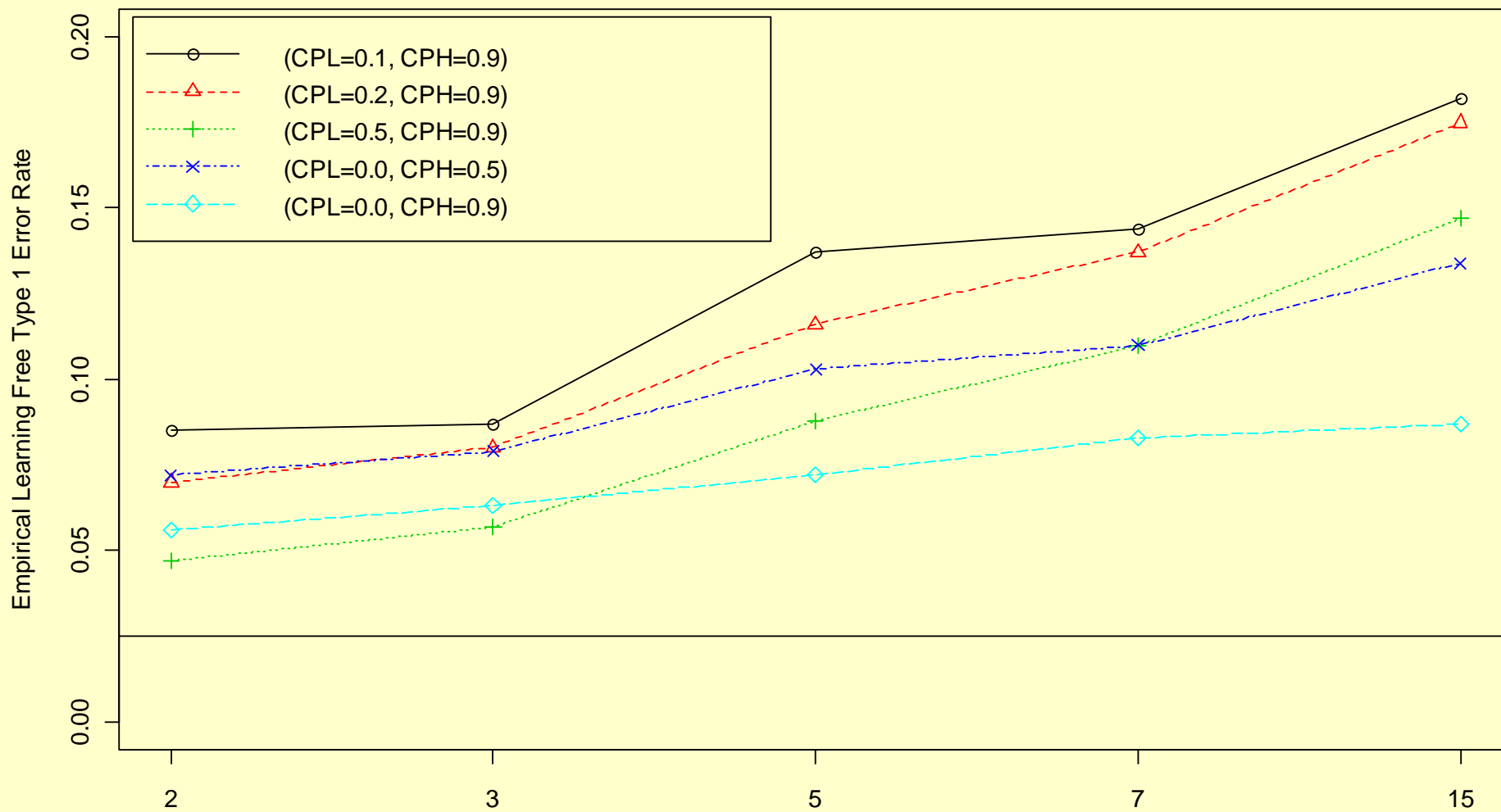
When learning data is used to select treatment arms in GSD setting

Select Dose Regimens Based on CP

For example:

- if no dose reaches $CP \geq cpH$; then final analysis tested at $\alpha / \#$ of kept doses
- if all doses have $CP < cpL$, then futility; else if at least one dose has $CP \geq cpH$, then select the lowest dose and final analysis tested at α

Figure 3. Dose Selection Based on Conditional Power of Clinical Endpoint at 50% Information Time

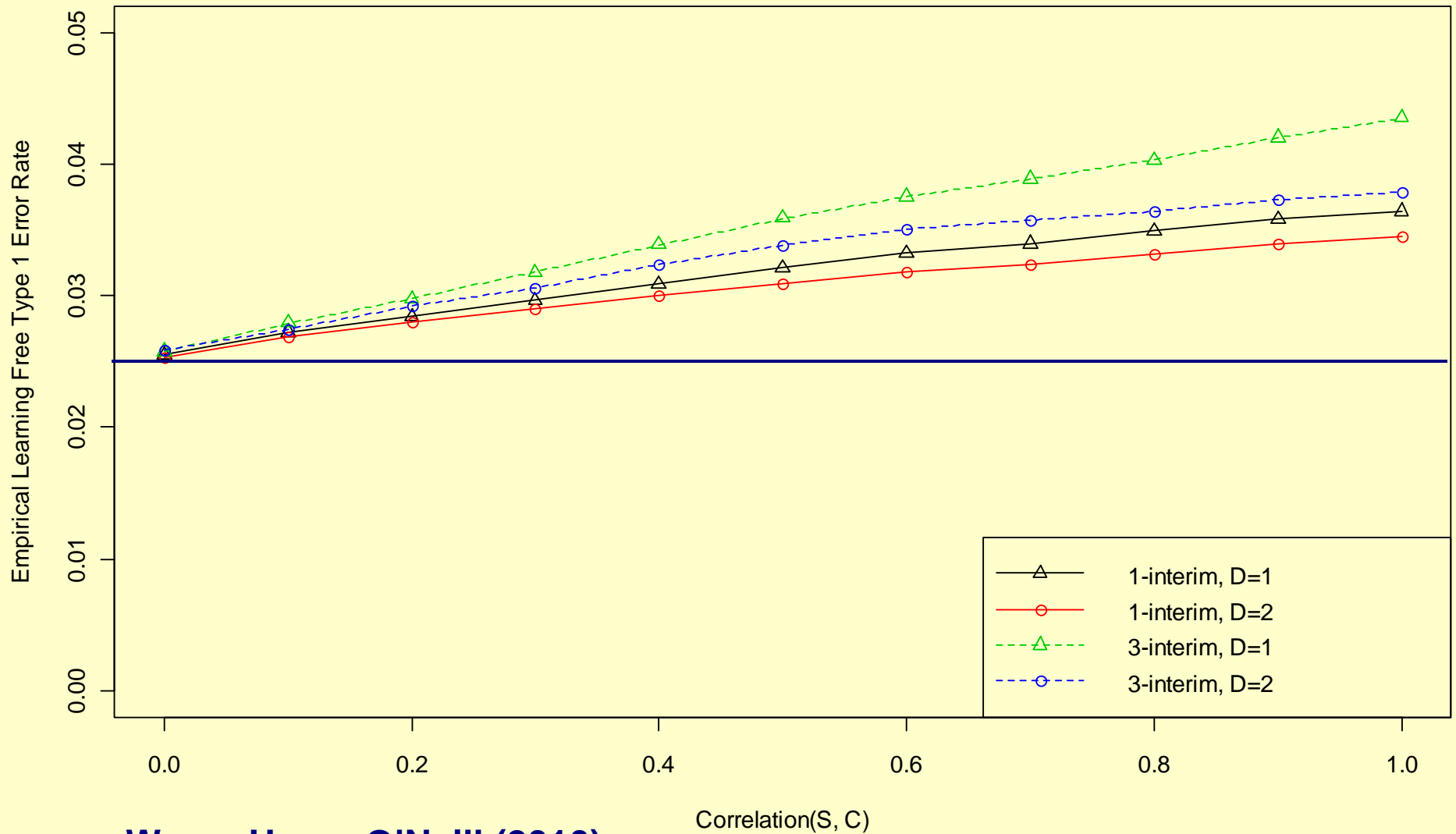


Wang, Hung, O'Neill (2010)

of Treatment Arms

Select between two treatment arms
based on surrogate endpoint,
final testing based on clinical endpoint

Figure 4. Dose Selection Based on Efficacy Biomarker Endpoint ($\Delta B=0$, $\Delta C=0$)



Wang, Hung, O'Neill (2010)

Learn/Confirm to select treatment arm(s)

- When dropping arm(s) occur, global H_0 changes;
Change in global H_0 is thought to be the flexibility of this approach; However, simply test remaining individual nulls at the seemingly controlled α -level, the learning-free type I error rate is unlikely controlled
- Important to weigh among when, what, how to do planned learning, such that statistical risk/benefit, viz., $\alpha, \beta : (1-\beta)$, is balanced

Learn/Confirm Endpoint Selection

Always start with the most interested endpoint;
consider interim switching based on pre-specified
interim adaptation rule

Timing of the selection

$$\tau = 1/6, 2/6, 3/6, 4/6, 5/6$$

Type I error rate spending

- .O'Brien-Fleming alpha spending rule (α_{OF})
- .Pocock alpha spending rule (α_{PK})
- .Final Analysis Only (α_{END})

Table 1. Fold increase in type I error rate associated with adaptive selection of one out of three pre-specified endpoints

| Correlation | $\tau = 1/6$ | | | $\tau = 2/6$ | | | $\tau = 4/6$ | | |
|-------------|---------------|---------------|----------------|---------------|---------------|----------------|---------------|---------------|----------------|
| | α_{OF} | α_{PK} | α_{END} | α_{OF} | α_{PK} | α_{END} | α_{OF} | α_{PK} | α_{END} |
| 0.0 | 1.9 | 2.0 | 1.9 | 2.2 | 2.0 | 2.2 | 2.5 | 1.8 | 2.6 |
| 0.3 | 1.7 | 1.8 | 1.7 | 2.0 | 1.8 | 1.9 | 2.2 | 1.6 | 2.2 |
| 0.5 | 1.6 | 1.6 | 1.5 | 1.8 | 1.6 | 1.7 | 1.9 | 1.5 | 1.9 |
| 0.9 | 1.2 | 1.2 | 1.1 | 1.2 | 1.1 | 1.2 | 1.1 | 1.1 | 1.1 |

Wang, Hung, O'Neill (2010)

Learn/Confirm Patient Subset

Molecularly untargeted \rightarrow Targeted*

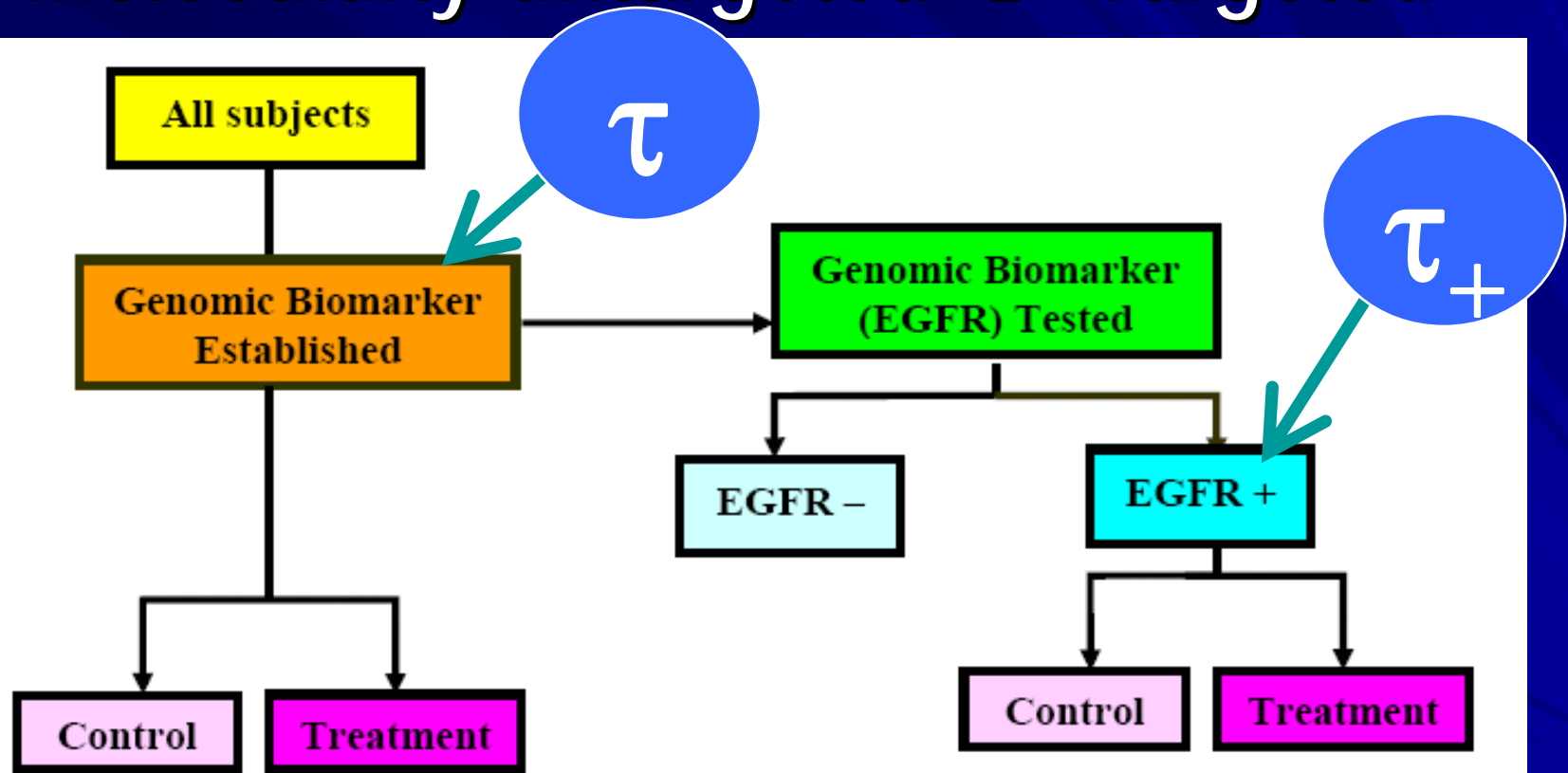


Figure. Use of Genomic Biomarker in Two-Arm Pharmacogenomics Clinical Trial

Wang, Hung (2005); Simon, Wang (2006); Wang, O'Neill, Hung (2007); Wang (2007)

Learning Free Type I Error Rate in Learn-and-Confirm

When early stage accumulating data is repeatedly used to learn one or more design parameters & also used with later data to confirm treatment effect:

Q: Can one argue that although the data is repeatedly used from what was learned & from within the same trial, but, because no formal statistical testing (relative to control) is performed, so, there is no need to spend α for repeated learning and essentially test, e.g., the best dose at 1-sided 2.5% at study end ?

Approaches for Dealing with Learning

(i) Conventional GSD that allows minimum changes (n, duration)

Learning on accumulated data via α -spending making it inferential

(ii) Combine Learning with Confirming (change global H_0 hypothesis)

(iii) Separate Confirming from Learning: Independent Data to Confirm



What Does Adaptive Design Offer ?

- ◆ Flexibility in planning can make it seamless
- ◆ Requires even more careful planning in early stages as scenario planning to increase chance of success in late stage registration trials
- ◆ Multiplicity in 2-stage: global H_0 & final H_0
- ◆ Can chance of late stage success with less prior data within trial actually be improved ?
- ◆ When more prior data external to the AD trial

Concluding Remarks

Intuition: the learning data in stage 1 should not be subject to type I error scrutiny if there is no formal interim analysis performed and only an adaptive selection of design parameters is made at stage 1

This research work shows that the intuition can be very misleading

Inappropriate use of Learn-and-Confirm strategy should not be overlooked



Encourage comments to the docket on aspects/issues you wish to bring to our attention

WWW.REGULATIONS.GOV

DOCKET # FDA-2010-D-0090

Or you can access via

www.fda.gov

and click on the link for

Dockets Management