Gatekeeping strategies in Phase III clinical trials with multiple endpoints and doses

Alex Dmitrienko (Quintiles, Inc)
Olga Marchenko (Quintiles, Inc)

EMA workshop
on multiplicity issues in clinical trials
Nov 2012

Outline

Multiple “sources” of multiplicity in clinical trials

Multiple endpoints and multiple dose-control comparisons

Gatekeeping procedures

Methods for building gatekeeping procedures

Development of gatekeeping procedures

Multiple “sources” of multiplicity

Multiple families of null hypotheses

Family 1	Primary hypotheses
Family 2	Secondary hypotheses
Family 3	Tertiary hypotheses

. . .

Multiplicity problems with a hierarchical structure

Case study

Latuda (lurasidone) Phase III program in patients with schizophrenia

Multiple doses

Two or three doses versus placebo

Multiple endpoints

Primary endpoint: Positive and Negative Syndrome Scale (PANSS) total score at Week 6

Secondary endpoints: Clinical Global Impression-Severity (CGI-S) score at Week 6 and PANSS total score at Day 4

Case study

Multiple objectives

Multiple doses: Improve success probability

Multiple endpoints: Strengthen lurasidone product label and create differentiating factors

Gatekeeping strategy

Powerful gatekeeping procedures were developed (Brechenmacher, Xu, Dmitrienko, Tamhane, 2011)

Importance of gatekeeping procedures was recognized in clinical publication (Meltzer et al., 2011)

Other examples

Osteoarthritis program

Two dose-placebo comparisons and three endpoints (WOMAC subscale scores, PGA)

Rheumatoid arthritis program

Two dose-placebo comparisons and four endpoints (DAS-28, ACR-20, HAQ, Sharp score)

Case study

Lurasidone Phase III trial

Multiple doses

Two doses versus placebo (Dose L, 40 mg/day; Dose H, 120 mg/day)

Multiple endpoints

Primary endpoint E1 (PANSS at Week 6)

Secondary endpoint E2 (CGI-S at Week 6)

Case study

Null hypotheses

	Dose L vs P	Dose H vs P
Endpoint E1	H_1	H_2
Endpoint E2	H_3	H_4

Overall Type I error rate (global familywise error rate) is controlled at two-sided $\alpha = 0.05$

Gatekeeping procedures

Definition

Multiple testing procedures for multiple families of null hypotheses

Type I error rate

Control Type I error rate over multiple families

Power

Optimal distribution of power by accounting for hierarchical structure of multiple families, e.g., more power for more important tests

Gatekeeping procedures

Main classes of gatekeeping procedures

Basic gatekeeping procedures based on Bonferroni test (Bretz et al., 2009; Burman et al., 2009)

Multistage gatekeeping procedures based on Bonferroni and more powerful tests (Dmitrienko, Tamhane and Wiens, 2008)

General mixture/gatekeeping procedures based on Bonferroni and more powerful tests (Dmitrienko and Tamhane, 2011; Kordzakhia and Dmitrienko, 2012)

Development of gatekeeping procedures

Principles

A. Incorporate all logical relationships among null hypotheses

B. Utilize available distributional information (joint distribution of hypothesis test statistics)

C. Select an optimal procedure (based on a relevant criterion under trial-specific assumptions)

Case study

Clinical information

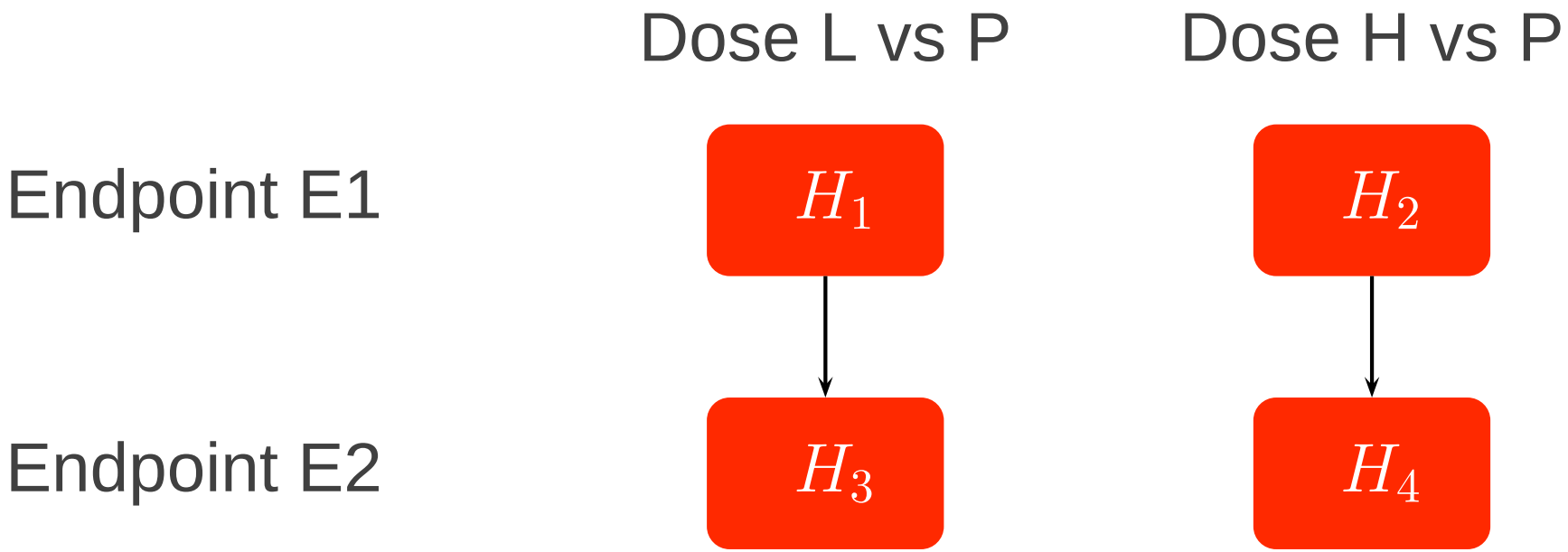
Establish efficacy based first on Endpoint E1 and then on Endpoint E2

Sufficient to establish efficacy for a single dose but highly desirable to demonstrate efficacy at both dose levels

No evidence of a positive dose-response relationship

A: Logical relationships

Clinical information

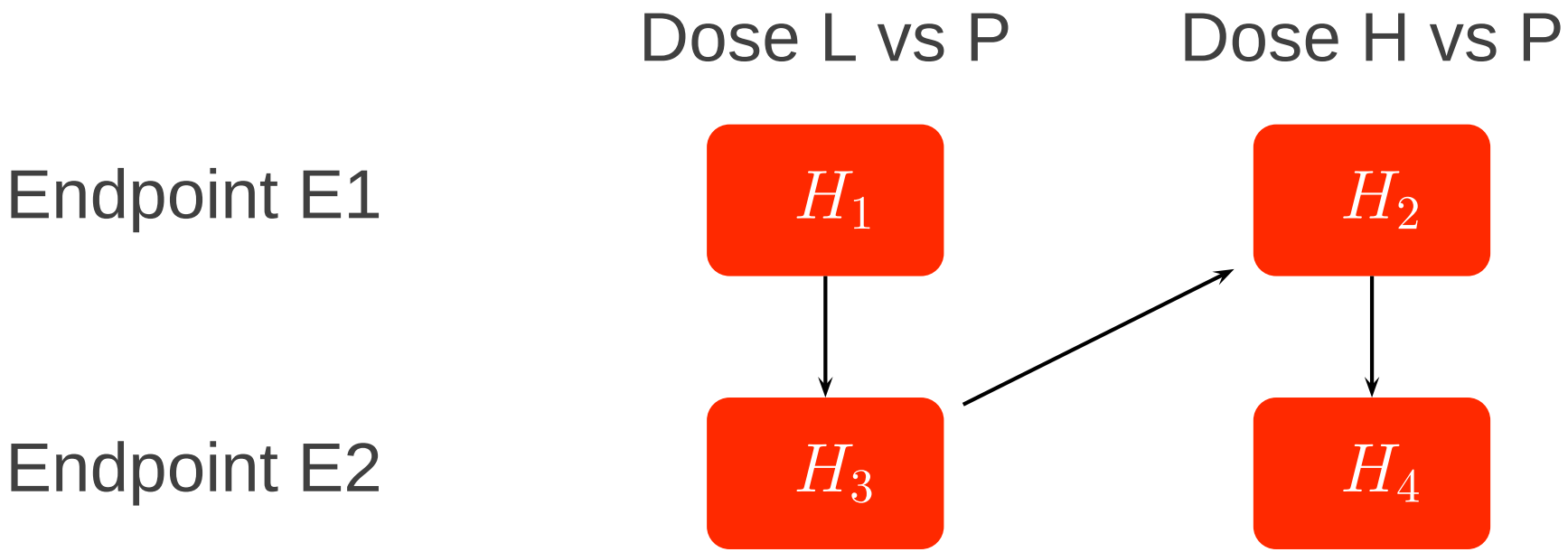


Take clinical information into account:

- H_3 depends on H_1
- H_4 depends on H_2

A: Logical relationships

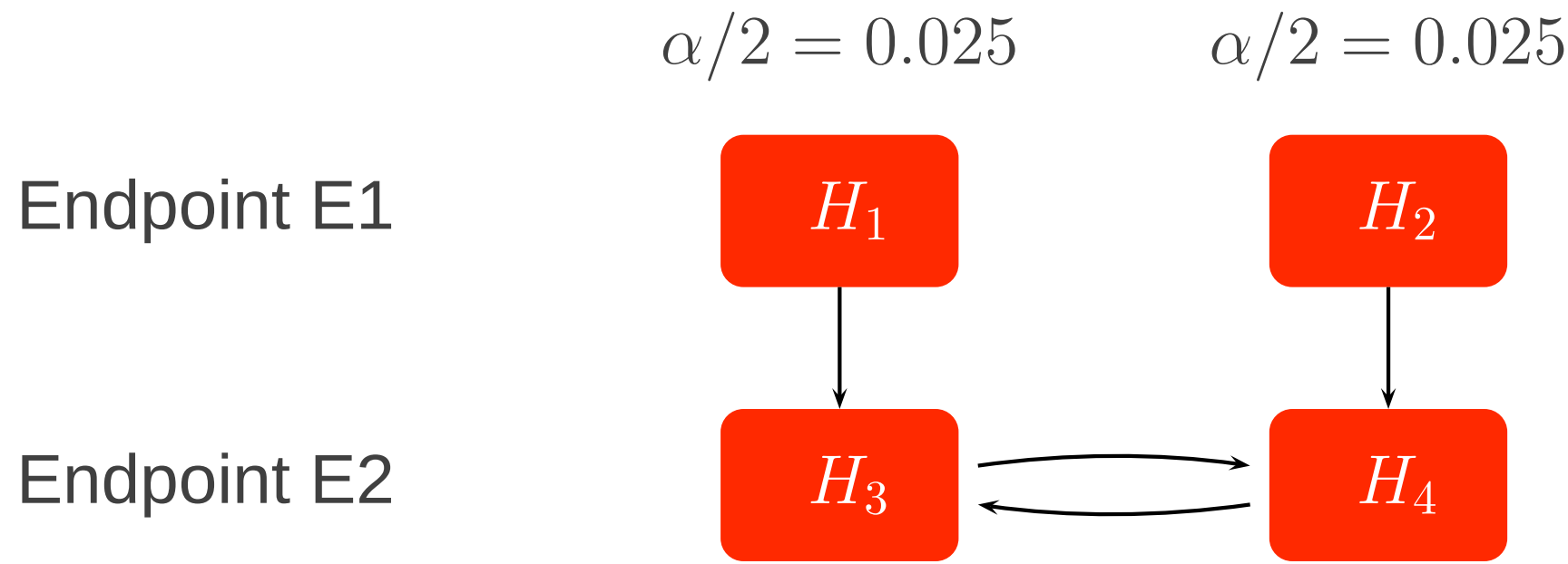
Serial testing strategy



Inflexible strategy which is not consistent with clinical objectives: H_2 and H_3 cannot be tested if H_3 is not rejected (Hung and Wang, 2009)

B: Distributional information

Gatekeeping procedure 1



Gatekeeping procedure based on an α-splitting method:

- α is split between H_1 and H_2
- α can be transferred between H_3 and H_4

B: Distributional information

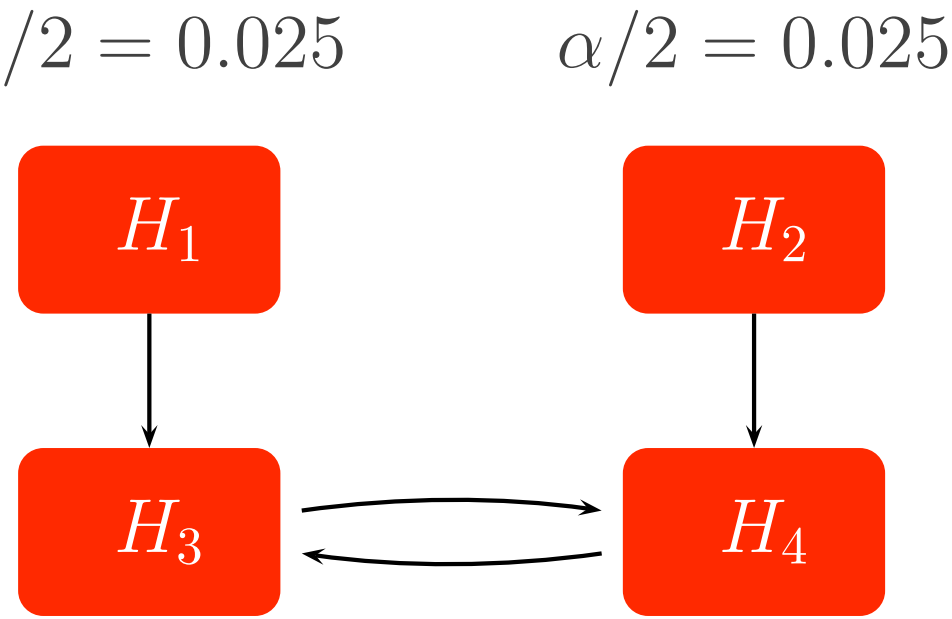
Gatekeeping procedure 1

Endpoint E1

Bonferroni

Endpoint E2

Holm



Bonferroni and Holm tests do not use available distributional information (test statistics within Families 1 and 2 are strongly positively correlated)

B: Distributional information

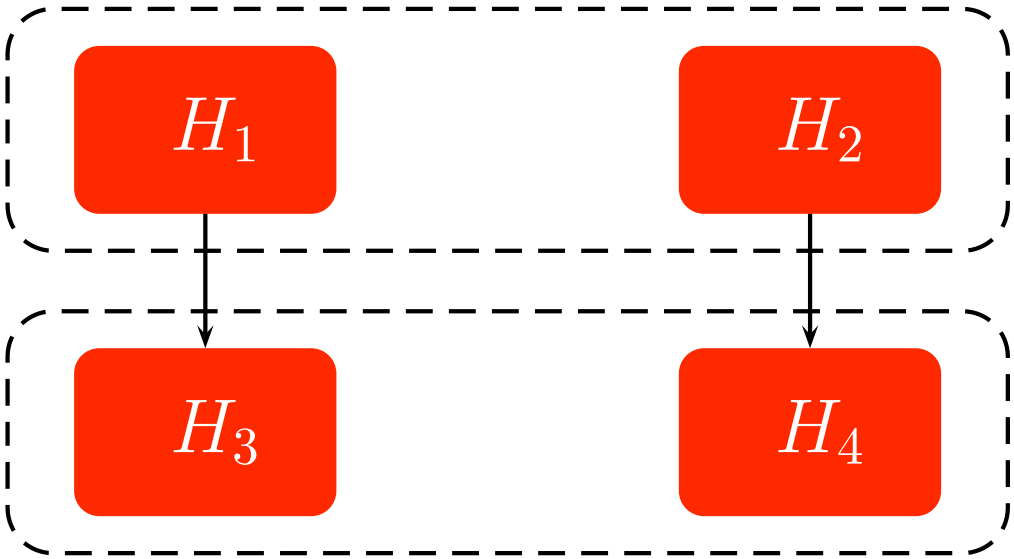
Gatekeeping procedure 2

Family 1

Powerful test

Family 2

Powerful test



Select tests that utilize available distributional information

Specify α propagation rules: how much error rate is transferred from Family 1 to Family 2

B: Distributional information

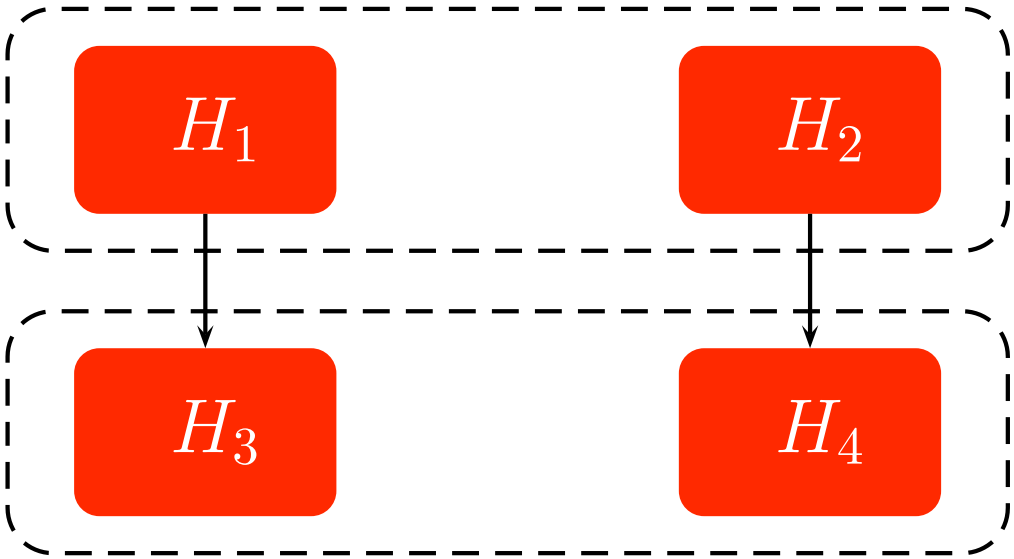
Gatekeeping procedure 2

Family 1

Truncated Hochberg

Family 2

Regular Hochberg



Mixture-based gatekeeping procedure:

- Truncated Hochberg test in Family 1 to enable flexible α propagation
- Regular Hochberg test in Family 2

C: Performance

Compare operating characteristics of candidate gatekeeping procedures

Gatekeeping procedure 1

Family 1: Bonferroni test

Family 2: Holm test

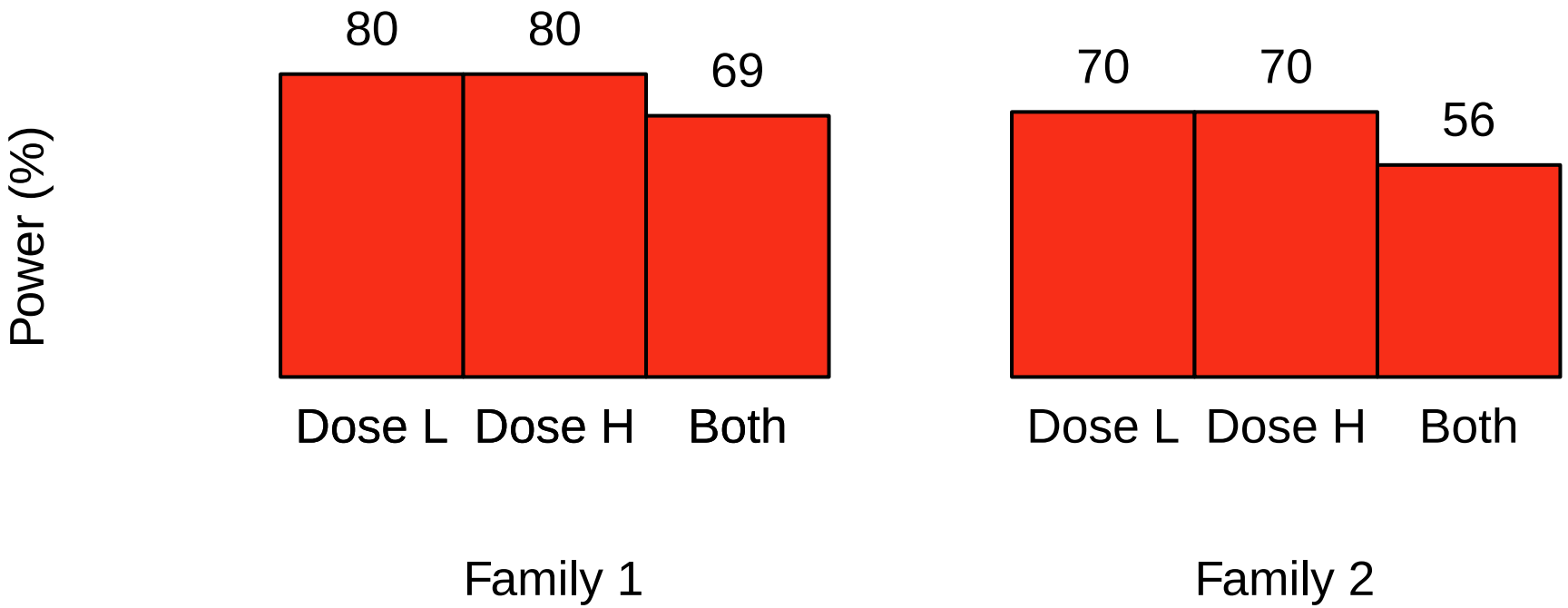
Gatekeeping procedure 2

Family 1: Truncated Hochberg test with truncation parameter of 0.7

Family 2: Hochberg test

C: Performance

Assumptions



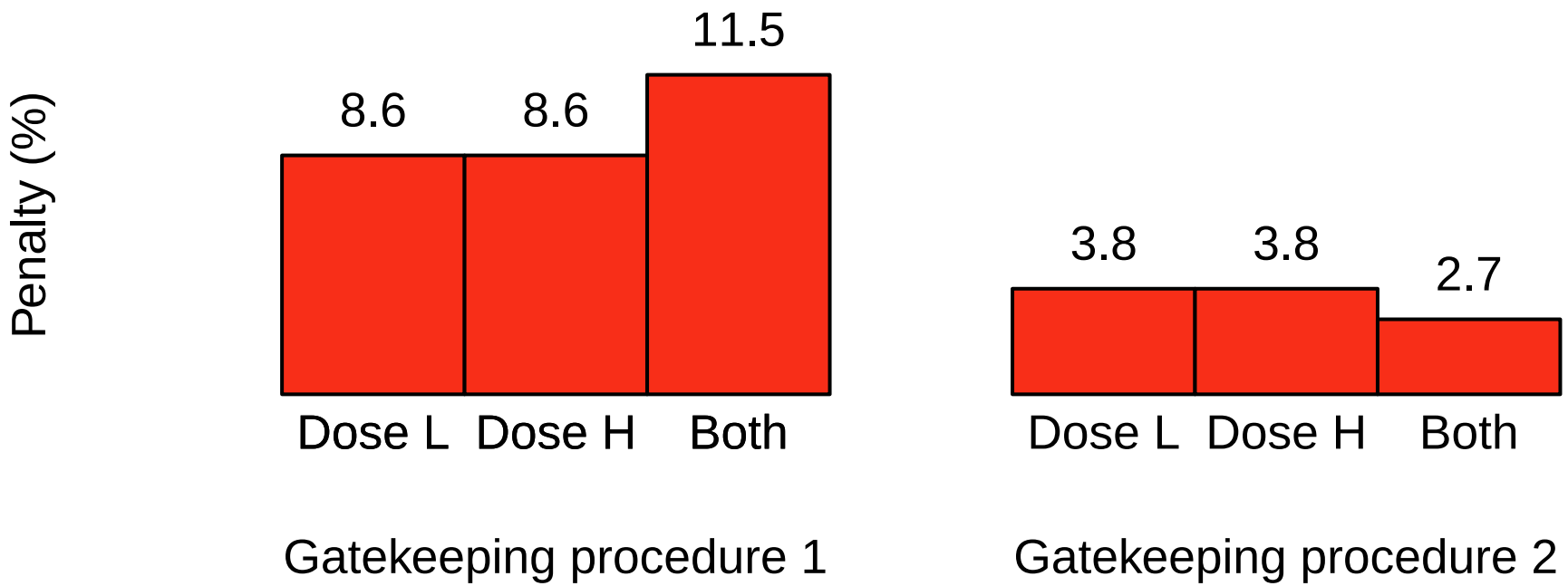
Dose L: Probability of achieving significant at Dose L

Dose H: Probability of achieving significant at Dose H

Both: Probability of achieving significant at both doses

C: Performance

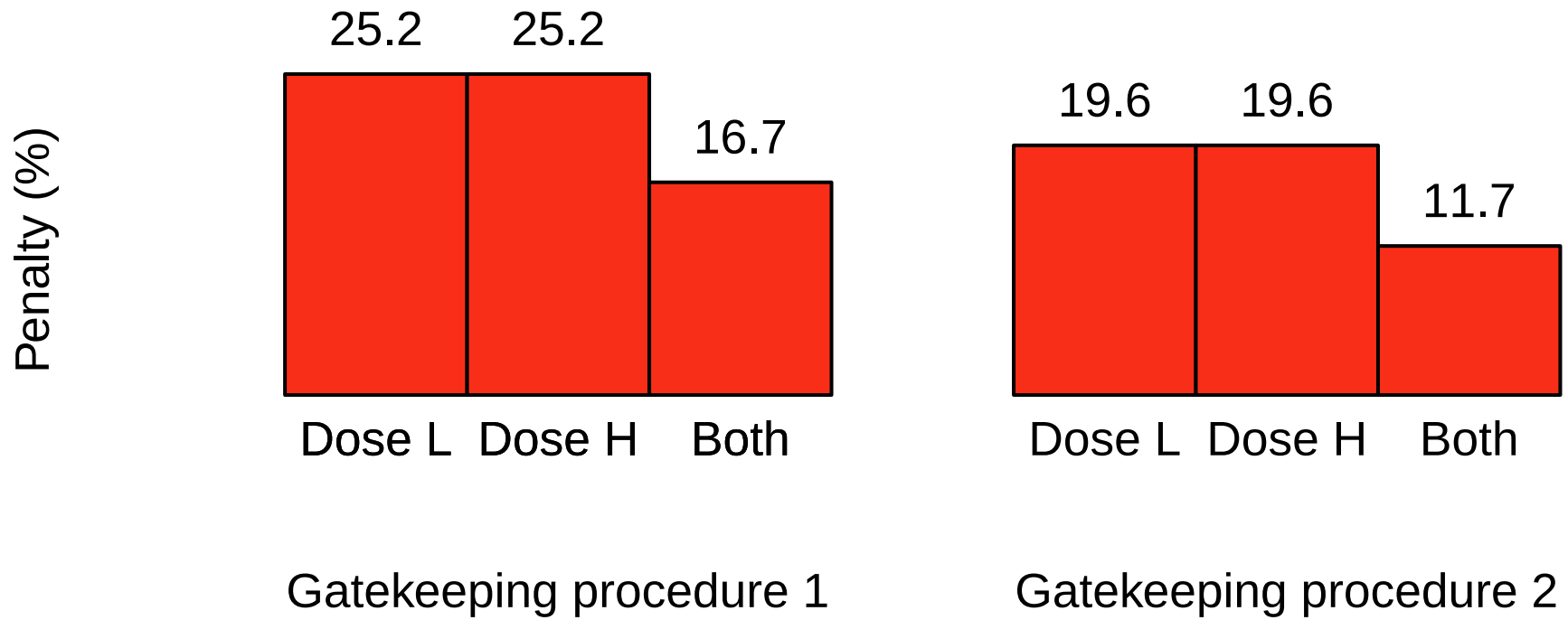
Multiplicity penalties in Family 1



Multiplicity penalty = Power before multiplicity adjustment − Power after multiplicity adjustment

C: Performance

Multiplicity penalties in Family 2



Multiplicity penalty = Power before multiplicity adjustment − Power after multiplicity adjustment

C: Performance

General evaluation criteria

Simple disjunctive power (one or more null hypotheses are rejected) or simple conjunctive power (all null hypotheses are rejected)

Subset disjunctive power (one or more null hypotheses are rejected in each family)

Weighted power

See Bretz, Maurer and Hommel (2011), Dmitrienko et al. (2011) for more information

Case study

Hochberg-based gatekeeping procedure

Endpoint	Dose	Raw p	Adjusted p
E1	L	0.001	0.002
	H	0.011	0.022
E2	L	0.006	0.011
	H	0.040	0.040

Both dose-placebo comparisons for Endpoints E1 and E2 are significant at $\alpha = 0.05$

Gatekeeping procedures in confirmatory trials

Type I error rate considerations

Control global error rate over multiple families

Power considerations

Based on powerful multiple tests

Clinical trial applications

Widely used in clinical trials to enrich product labels and provide important clinical information to physicians and patients (lurasidone product label)

References

Brechenmacher T, Xu J, Dmitrienko A, Tamhane AC. (2011). A mixture gatekeeping procedure based on the Hommel test for clinical trial applications. *Journal of Biopharmaceutical Statistics*. **21**, 748–767.

Bretz F, Maurer W, Brannath W, Posch M. (2009). A graphical approach to sequentially rejective multiple test procedures. *Statistics in Medicine*. **28**, 586–604.

Burman CF, Sonesson C, Guilbaud O. (2009). A recycling framework for the construction of Bonferroni-based multiple tests. *Statistics in Medicine*. **28**, 739–761.

References

Bretz F, Maurer W, Hommel G. (2011). Test and power considerations for multiple endpoint analyses using sequentially rejective graphical procedures. *Statistics in Medicine* **30**, 1489–1501.

Dmitrienko A, Offen WW, Westfall PH. (2003). Gatekeeping strategies for clinical trials that do not require all primary effects to be significant. *Statistics in Medicine*. **22**, 2387–2400.

Dmitrienko A, Tamhane AC, Wiens BL. (2008). General multistage gatekeeping procedures. *Biometrical Journal*. **50**, 667–677.

References

Dmitrienko A, Tamhane AC. (2009). Gatekeeping procedures in clinical trials. *Multiple Testing Problems in Pharmaceutical Statistics*. Dmitrienko A, Tamhane AC, Bretz F. (editors). Chapman and Hall/CRC Press, New York.

Dmitrienko A, Tamhane AC. (2011). Mixtures of multiple testing procedures for gatekeeping applications in clinical trials. *Statistics in Medicine*. **30**, 1473–1488.

Dmitrienko A, Millen BA, Brechenmacher T, Paux G. (2011). Development of gatekeeping strategies in confirmatory clinical trials. *Biometrical Journal*. **53**, 875–893.

References

Hung J, Wang SJ. (2009). Some controversial multiple testing problems in regulatory applications. *Journal of Biopharmaceutical Statistics*. **19**, 1–11.

Kordzakhia G, Dmitrienko A. (2012). Superchain procedures in clinical trials with multiple objectives. *Statistics in Medicine*. In press.

Meltzer et al. (2011). Lurasidone in the treatment of schizophrenia: a randomized, double-blind, placebo- and olanzapine-controlled study. *American Journal of Psychiatry*. **168**, 957–967.