Testing Monotherapy and Combination Therapy in One Trial with Biomarker Consideration

ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop

Linda Sun, S. Peter Kang, Cong Chen

2019-09-25

MERCK

Outline

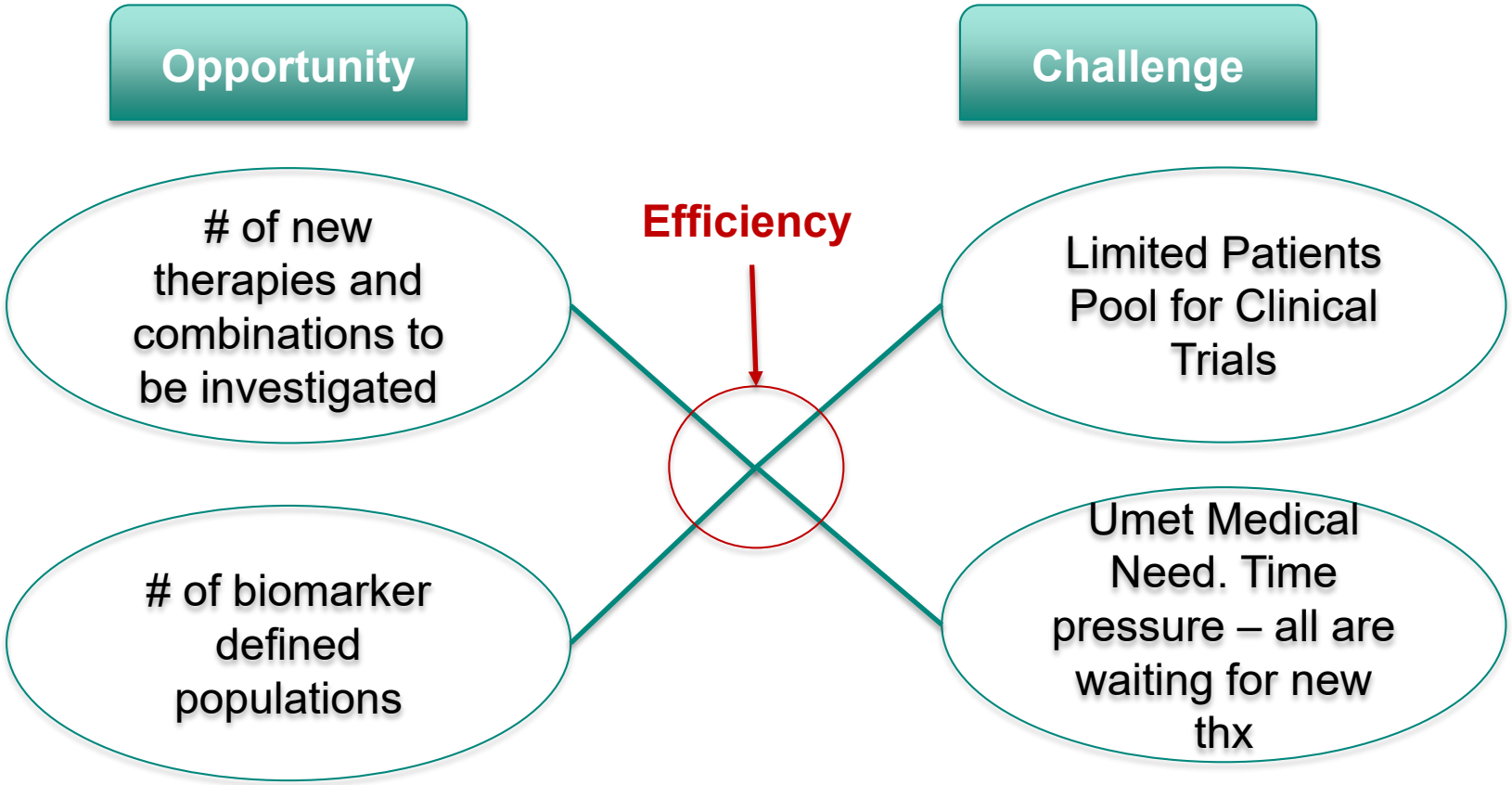
- Motivation
- Proposed Design
  - Analysis method
  - Multiplicity consideration
- Power and sample size
- Discussion

Motivation

MERCK

Called to be efficient

- Trial Design in Oncology Drug Development



MERCK

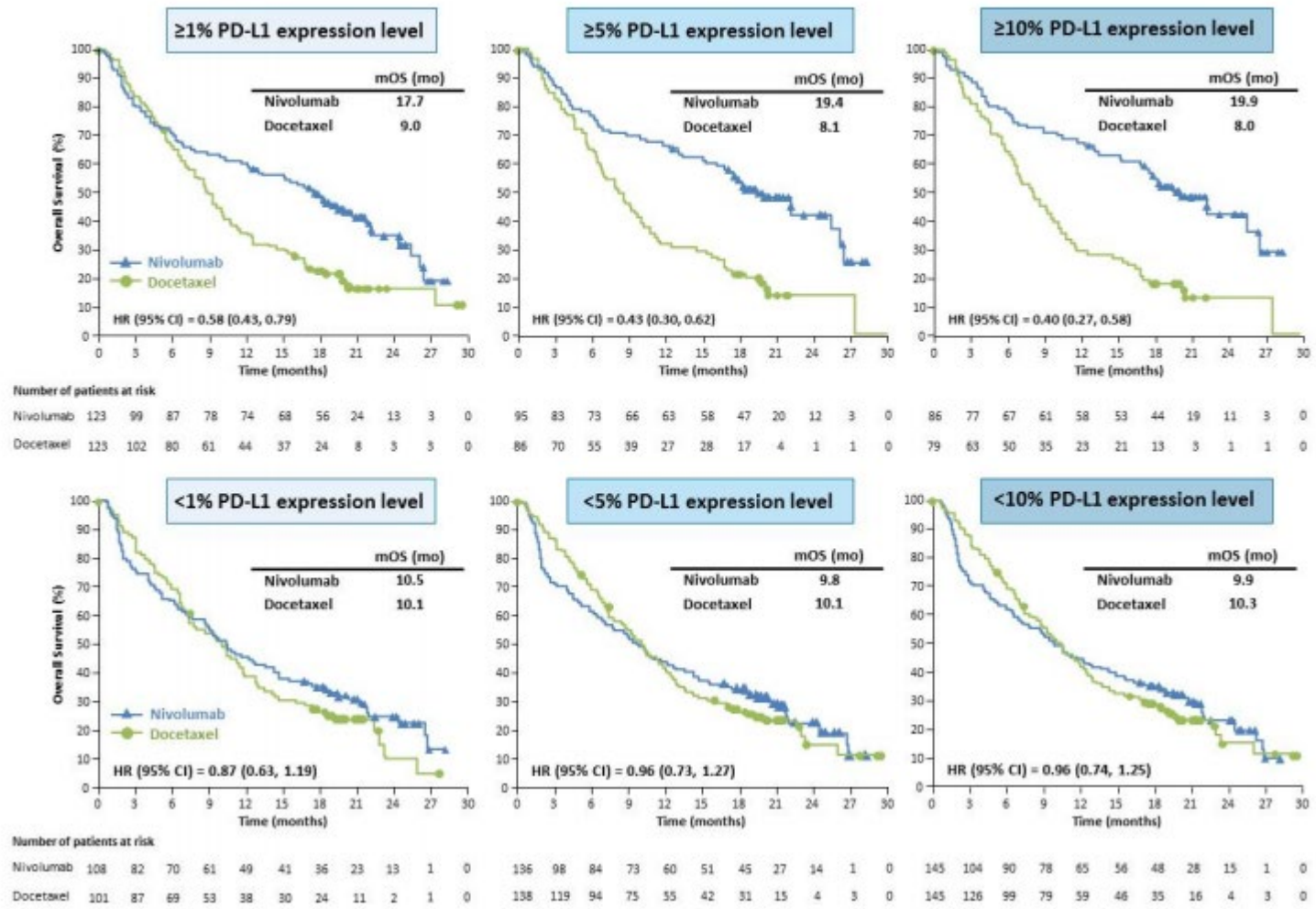
A common scenario in oncology drug development

- Monotherapy of experimental drug (E)
  - Treatment effect is bigger in BMX enriched population
  - Maybe superior to SOC in BMX+ population
  - Maybe comparable or even inferior to SOC in BMX- population
- Combination of E + SOC
  - Maybe superior to SOC in both BMX+ and BMX-, i.e. all-comer population

MERCK

Example – PD1 check point inhibitor in NSCLC

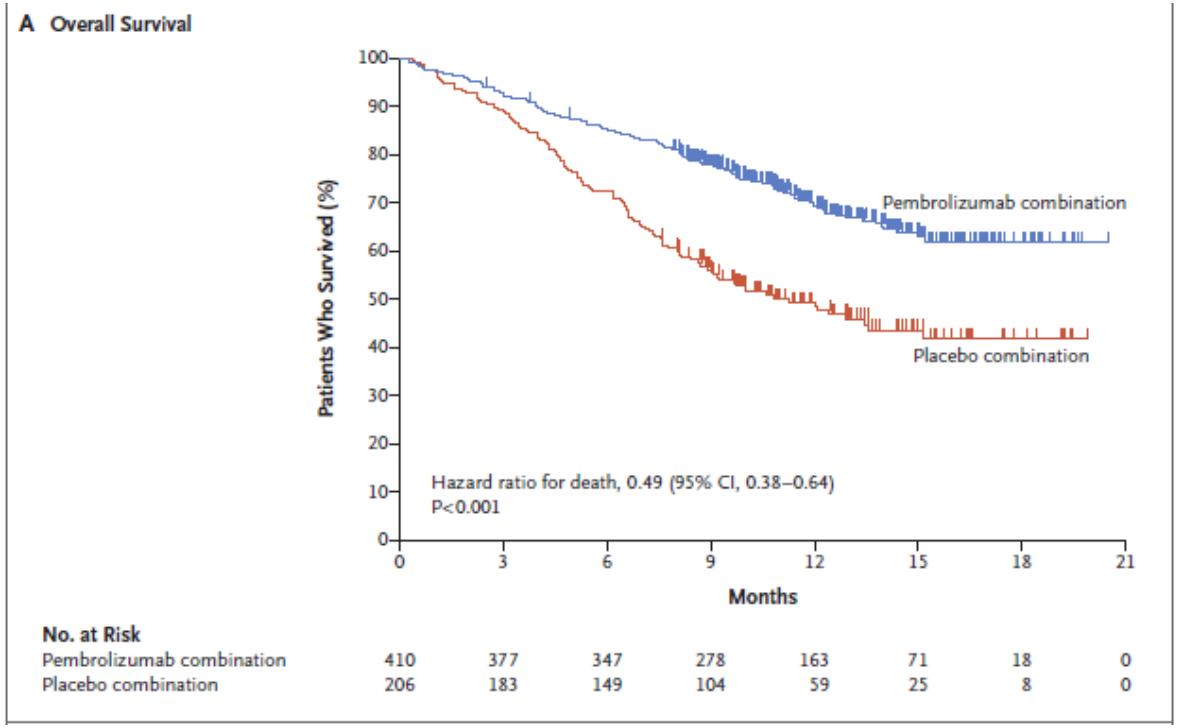
Figure S8B. Kaplan-Meier Plots of Overall Survival With Extended Follow up at the 1%, 5% and 10% PD-L1 Expression Levels.



Monotherapy treatment effect by PD-L1 level in 2L NSCLC

- Nivolumab, Checkmate 057, NEJM 2015

Example – PD1 check point inhibitor in NSCLC



Combination therapy is superior to SOC in all- comers and all PD-L1 subgroups

- Pembrolizumab, Keynote-189, NEJM 2018

Example – PD1 check point inhibitor in NSCLC

- For monotherapy
  - Even though it's comparable to 2L chemo in in PD-L1 - patients, there was concern that it may not be able to measure up to 1L chemo
  - Therefore, testing monotherapy in 1L PD-L1- patients may expose patients to potentially inferior treatment
  - Monotherapy PD1 is to be tested in BMX+ enriched 1L NSCLC
- For combination
  - Since the experimental treatment is add-on, there is no concern to test the combination thx in all-comer population

MERCK

Motivation

- Notice: the same SOC on the control arm for mono and combo trials, why not share control in one trial?
- Possible reasons for the two-trial strategy
  - The dose for combination therapy is determined later than the monotherapy dose
  - The populations are different

Proposed design provides an solution

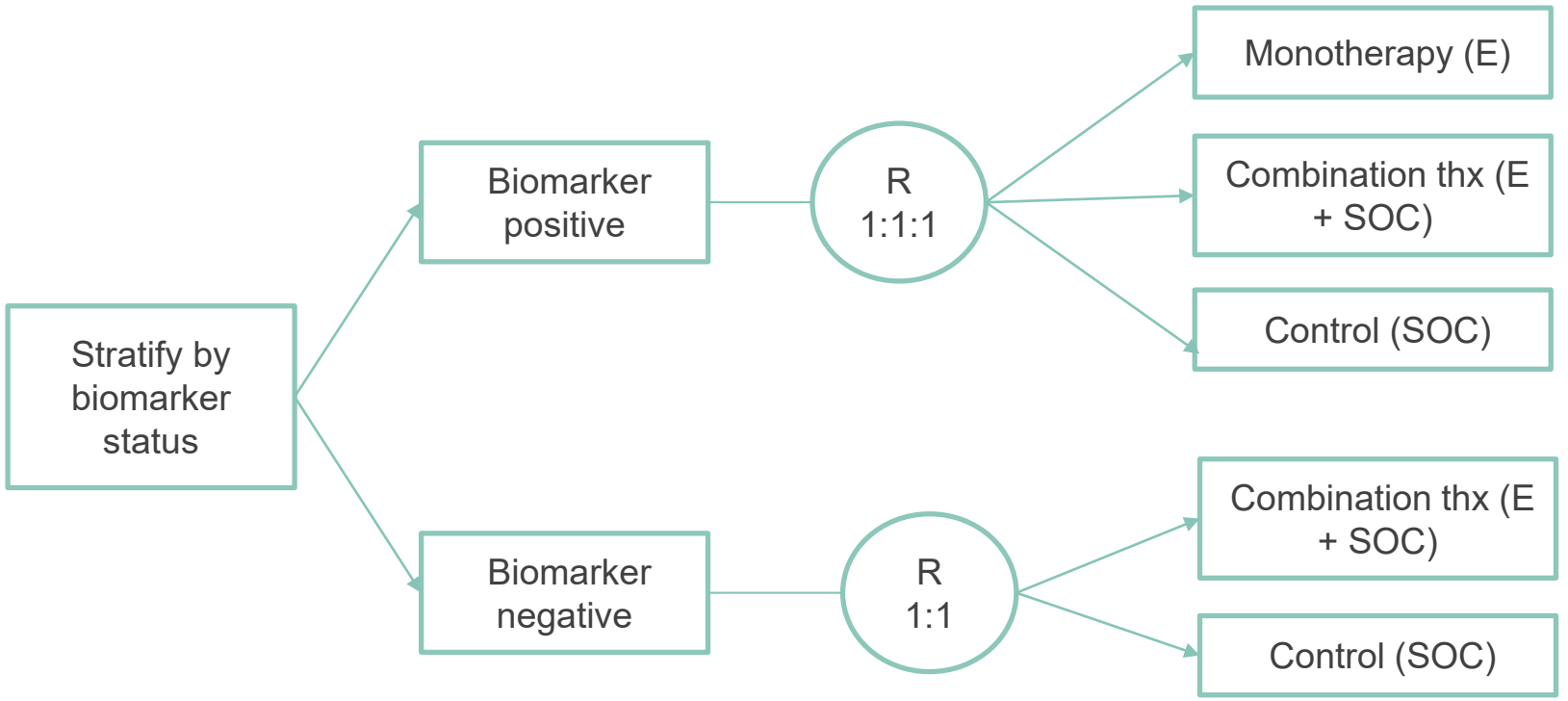
No longer an issue for follow on tumor types

MERCK

Proposed Design

MERCK

Study Design



Hypotheses (Primary endpoint: OS)

H1: Monotherapy vs. SOC in BMX+ population

H2: Combination thx vs. SOC in all-comer population

Analysis Method

- For H1 (monotherapy)
  - Regular log-rank test and Cox-regression method
- For H2 (combination therapy)
  - Population is skewed – BMX+ prevalence is lower than the prevalence in true all-comer
  - Two-step log-rank test and Cox-regression method
    - Step 1: analyze BMX+ and BMX- separately
    - Step 2: weighed sum to combine.
  - Weight of BMX+ patients vs BMX- patients = 3:2, pre-determined by randomization ratio and no estimation
  - Minor loss of efficiency with two-step log-rank test

MERCK

Multiplicity Consideration

- The proposed design is a multi-arm design with shared control group.
- Is it necessary to adjust for multiple comparison (experimental arms vs. control) in the multi-arm single trial when no such adjustment would be required had these comparisons been made in separate two-arm trials?
- Key question: Were the experimental arms put together in a single trial because the corresponding research questions are related, or primarily, because of efficiency and logistical reasons?
- No multiplicity adjustment is needed in multi-arm trials that are designed for logistical efficiency. (Freidlin 2008, Howard 2018)
- (Same thinking for multiplicity consideration in Umbrella/Platform studies)

MERCK

A distinction

- Distinction between factorial design and the proposed design
  - Factorial design
    - A+B, A, B, vs. SOC
    - Arm A and B are in the trial for component contribution
  - The proposed design
    - **H1**: E vs. SOC in BMX+ population
    - **H2**: E+SOC vs. SOC in all-comer population
    - The monotherapy arm (E) in the proposed design doesn't serve as component contribution
    - Arm E and E+SOC in one trial are purely for logistical efficiency
- No multiplicity adjustment
  - If H1 and H2 are tested in two separate trials, each with alpha = 0.025 (one-sided), then in the proposed design, H1 and H2 will also each be tested at 0.025 (one-sided) in the proposed design
  - The sharing of a common control actually reduces the FWER as compared to separate trials

Power and Sample Size

MERCK

Sample Size Comparison - The Hypothetical Example

	One-trial Design (Proposed Design)	Two-trial Design
Sample size for H1	326	326
Sample size for H2	489	472
Sample size saved due to shared control for H1/H2	(-) 61	0
Total sample size	754	798
# of screened patients who cannot be enrolled solely due to being BMX-	408	652

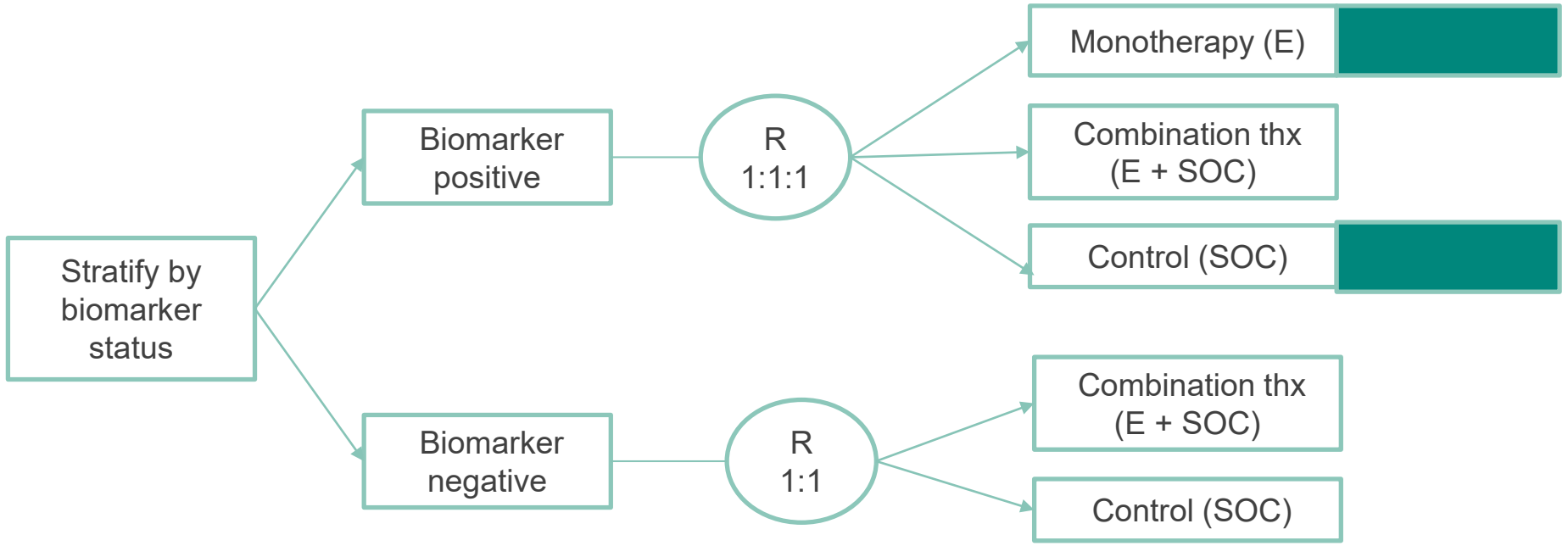
Biomarker prevalence = 1/3, HR (OS) mono vs SOC = 0.65, HR (OS) combo vs SOC = 0.70, one-sided alpha = 0.025, power = 90%, 70% randomized patients have events by the time of final analysis.

Trial cost per patient > $100K

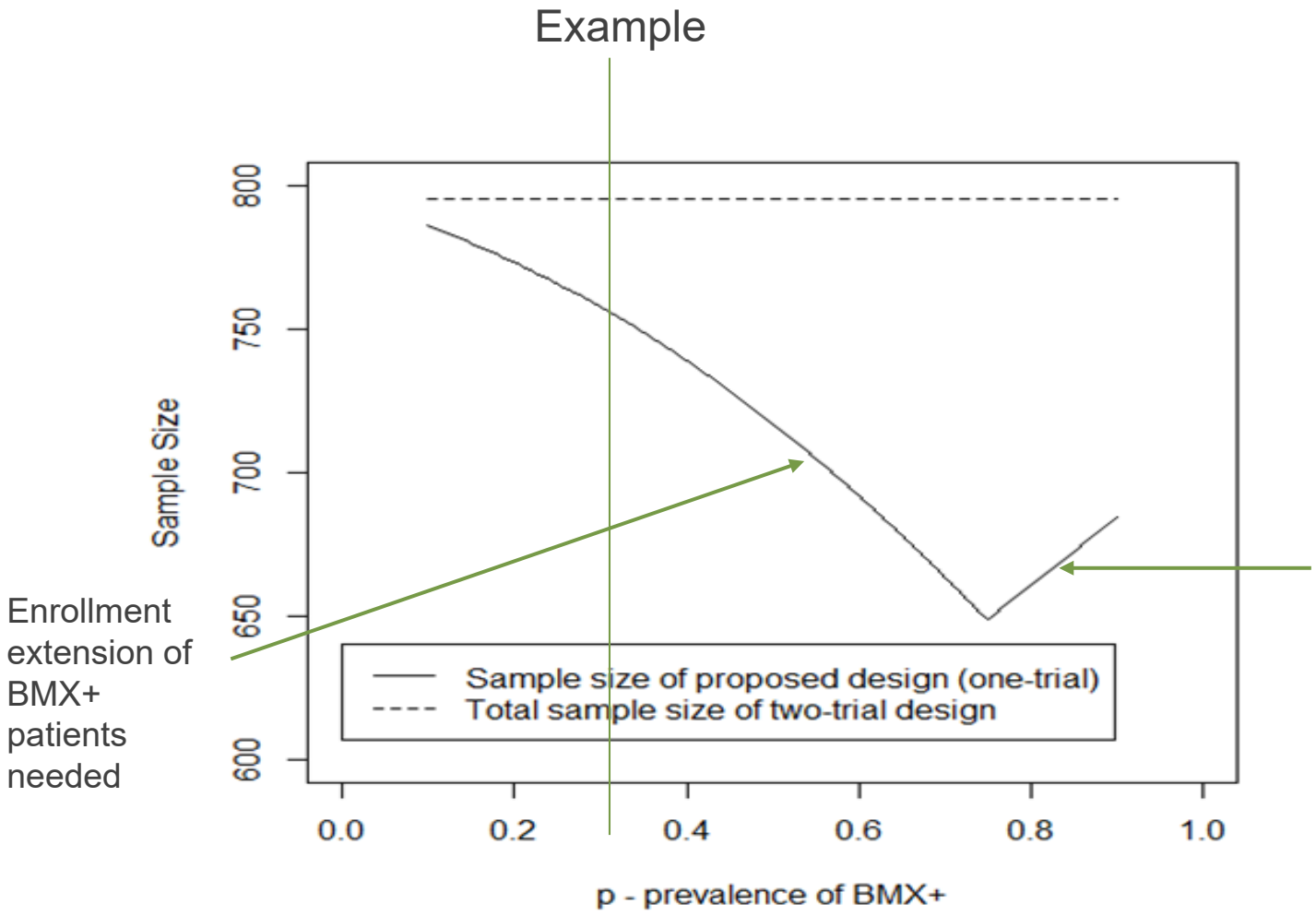
MERCK

How much sample size the proposed design can save?

- Depends on sample size for H1 relative to H2
  - Biomarker prevalence *p*
  - Target HR for H1 and H2
  - May need enrollment extension for H1



Impact of BMX+ Prevalence on Sample Size



- The BMX+ patients randomized to the control arm contributes to test both H1 and H2
- **In general, the higher the BMX+ prevalence, the more efficient the proposed design**
- However, when BMX+ prevalence is higher than a certain point, the sample size of the proposed design is driven by H2 only. For ease of operation, we don't cap the enrollment of BMX+. Thus the design loses some efficiency.

Discussion

MERCK

Logistical Advantages of the Proposed Design

- If the two-trial design are conducted sequentially, the proposed design can save time
- If the two-trial design enroll at the same time,
  - For BMX- patients, screen failure from Trial 1 cannot be automatically considered for Trial 2
  - For BMX+ patients, which trial to enroll? Trial 1 or Trial 2?
    - Due to toxicity consideration, investigators may enroll poor prognostic BMX+ patients to Trial 1

MERCK

Extensions of Design and Methodology

- Design extension
  - Add group sequential feature
  - Population adaptation:
    - Start with three-arm all-comer trial
    - Futility of BMX- patients for either monothx or combination, or both
- Two-step method
  - Umbrella platform, evaluating treatments targeting different biomarkers
  - How to assign patients with multiple biomarkers of interest
  - Population adjustment to recover true all-comer

Further Multiplicity Consideration

- If multiplicity adjustment is required for H1 and H2 in one-trial design (due to philosophical consideration), but not in two-trial design
  - The one-trial design requires more sample size when BMX+ prevalence is low (e.g. 33%)
  - The one-trial design may require the same sample size as the two-trial design when BMX+ prevalence is higher (e.g. 67%)

MERCK

Summary

- Cancer drug development
  - Fast-growing
  - Complex
- Proposed design is efficient
  - Sample size
  - # of patients screened
  - Time
- The use of the two-step approach can be extended to other areas

MERCK

References

1. Sun L, Kang P., Chen C. Testing of Monotherapy and Combination Therapy in One Trial with Biomarker Consideration, Contemporary Clinical Trials 2019, 82 (2019): 53-59.

2. Woodcock J, LaVange LM. Master protocols to study multiple therapies, multiple diseases, or both. N Engl J Med. 2017;377(1):62–70.

3. Parmar MK, Carpenter J, Sydes MR. More multiarm randomised trials of superiority are needed. Lancet 2014, 384:283–284.

4. Howard DR, Brown JM, Todd S, Gregory WM. Recommendations on multiple testing adjustment in multi-arm trials with a shared control group. Statistical Methods in Medical Research 2018, 27:1513-1530.

5. Proschan M and Follman D. Multiple comparisons with control in a single experiment versus Separate Experiments: Why Do We Feel Differently? Am Stat 1995, 49:144.

6. Freidlin B, Korn EL, Gray R, et al. Multi-arm clinical trials of new agents: some design considerations. Clin Cancer Res 2008; 14: 4368–4371

MERCK

Back Up

MERCK

Multiplicity

- Distinction between research questions and test statistics
  - If H1 and H2 are considered as separate research questions when tested in two trials, then the same when tested in one trial.
  - The test statistics are related due to shared control in the one-trial design. This correlated reduces FWER compared to two-trials design

$$FWER = \Pr(X_1 > Z_\alpha \text{ or } X_2 > Z_\alpha) = 2\alpha - \Pr(X_1 > Z_\alpha, X_2 > Z_\alpha) \leq 2\alpha - \alpha^2$$

MERCK

Impact of BMX+ Prevalence on Sample Size

- Total sample size
- When enrollment extension of BMX+ is needed

$$T = (1 - \frac{2}{3}p)(1 + \frac{1}{2}p)\frac{(Z_{1-\alpha}+Z_{1-\beta})^2}{\log(HR_2)^2} \cdot 4 + \frac{(Z_{1-\alpha}+Z_{1-\beta})^2}{\log(HR_1)^2} \cdot 4$$

- When enrollment extension is not needed

$$T = \left(1 + \frac{1}{2}p\right)\frac{(Z_{1-\alpha}+Z_{1-\beta})^2}{\log(HR_2)^2} \cdot 4$$

MERCK