


Recent Approaches to Surrogate Endpoint Validation

Stuart G. Baker, Sc.D.
National Cancer Institute
sb16i@nih.gov



What do we mean by validation of a marker or endpoint?

- A marker or endpoint is validated if there is confidence in proceeding to the *next stage of evaluation*
- Methodology depends on the application
 - Biomarkers for the early detection of cancer
 - Biomarkers for predicting cancer recurrence
 - Biomarkers for targeting intervention in treatment trials
 - Surrogate endpoints

Examples of validation

- Biomarkers for early detection of cancer
 - **Next stage:** further study as a trigger for early intervention with cancer-mortality endpoint
 - **Validation:** high sensitivity and specificity in an independent test sample
- Biomarkers for predicting cancer recurrence
 - **Next stage:** randomized trial of standard versus new treatment for those with poor prognosis
 - **Validation:** high predictive value positive and negative in an independent test sample

■ Examples of validation (continued)

- Biomarkers for targeting intervention in a treatment trial
 - **Next stage:** recommendation for population with the biomarker
 - **Validation:** randomized trial of intervention in subjects with the biomarker
- **Surrogate endpoints**
 - **Next stage:** recommendation for population (although sometimes further study)
 - **Validation:** correct conclusions about effect of intervention on true endpoint (in trial with both surrogate and true endpoints)



Surrogate endpoint: definition

Measure or indicator of a biological process that is

- (a) used to make conclusions about the effect of intervention on a true endpoint that is a health outcome
- (b) obtained sooner, at less cost, or less invasively than a true endpoint



Validating a surrogate endpoint

- **NEXT STAGE:** Application trial: surrogate but not true endpoint is observed
 - *What is the effect of intervention on true endpoint?*
- **VALIDATION:** Validation trial in which both surrogate and true endpoints are observed
 - *Are the conclusions about the effect of intervention on true endpoint the same when based on*
 - (i) only surrogate endpoint*
 - (ii) only the true endpoint ?*



Issues in validating a surrogate endpoint

- Hypothesis testing framework
- Estimation framework
- Caveats
- Recommendations

Hypothesis Testing Framework

Validation trial

Surrogate endpoints

True endpoints

- (1) Prentice Criteria IMPLIES
- (2) Rejecting null hypothesis for surrogate endpoint implies rejecting null hypothesis for true endpoint

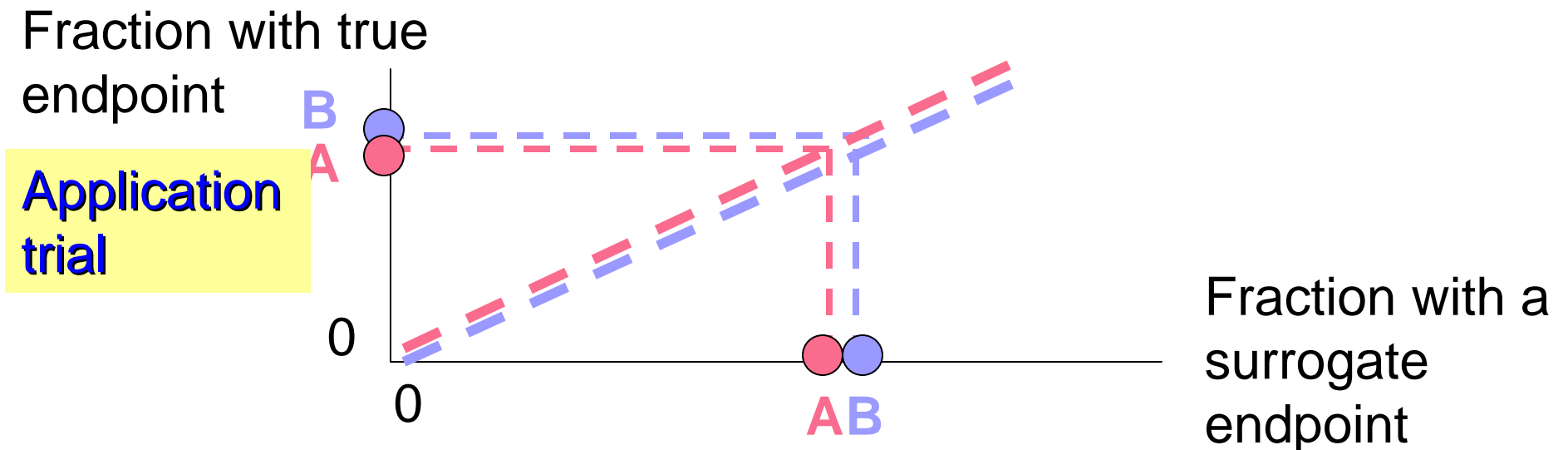
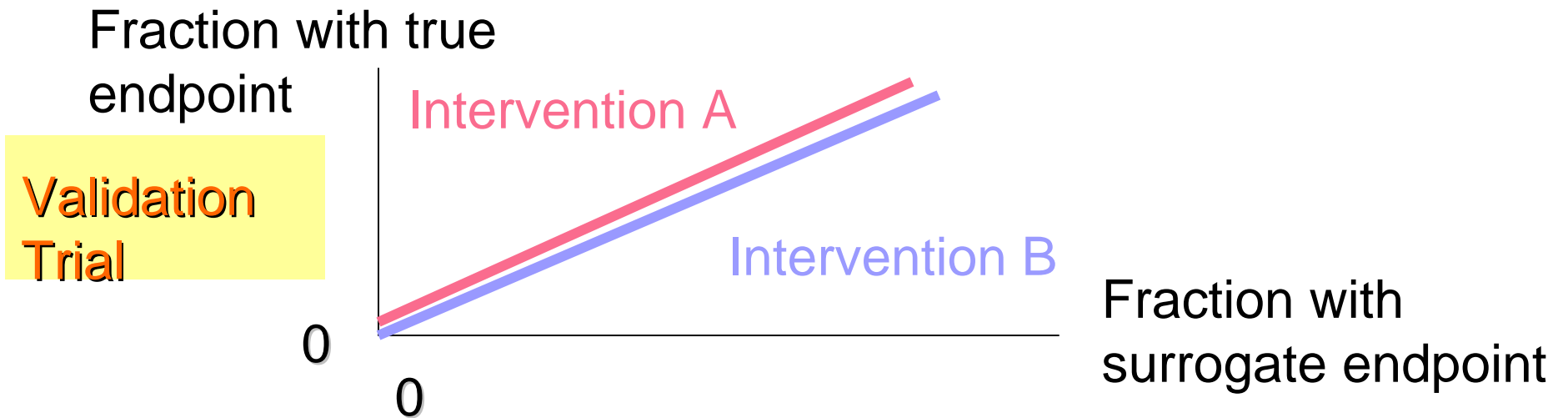
Validation

Application trial

Surrogate endpoint

Test null hypothesis using surrogate endpoint

Prentice Criterion “holds” (identical lines for A and B)



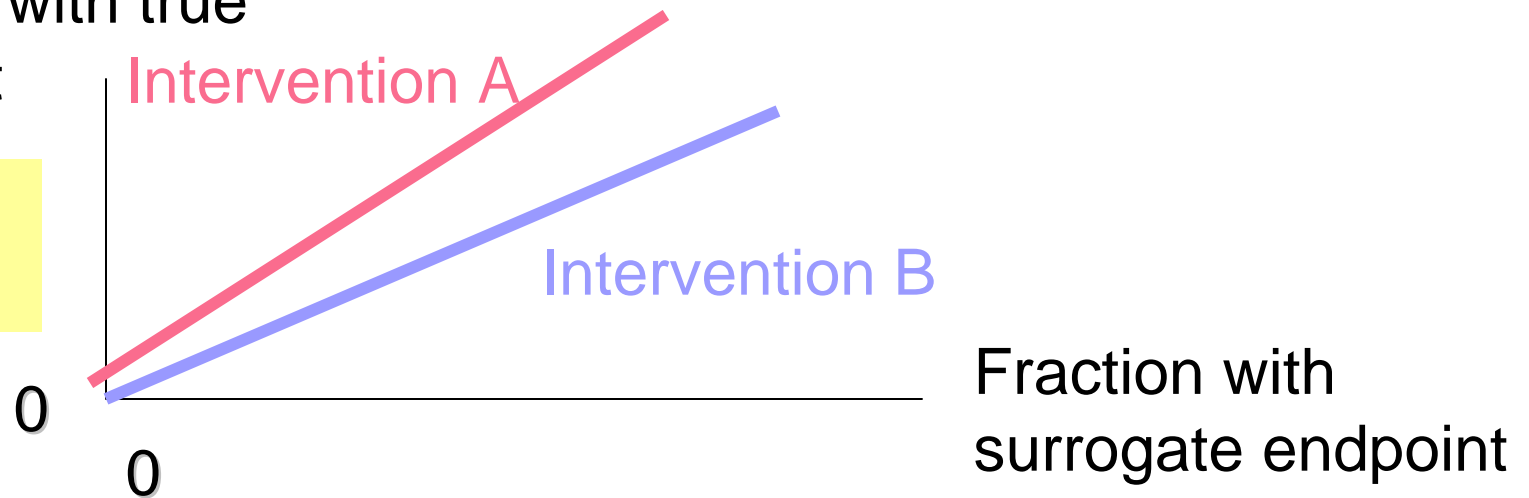
$H_0: A = B$ for true implies $H_0: A = B$ for surrogate

Reject null for surrogate implies reject null for true

■ Prentice Criterion does NOT hold (lines for **A** and **B** differ)

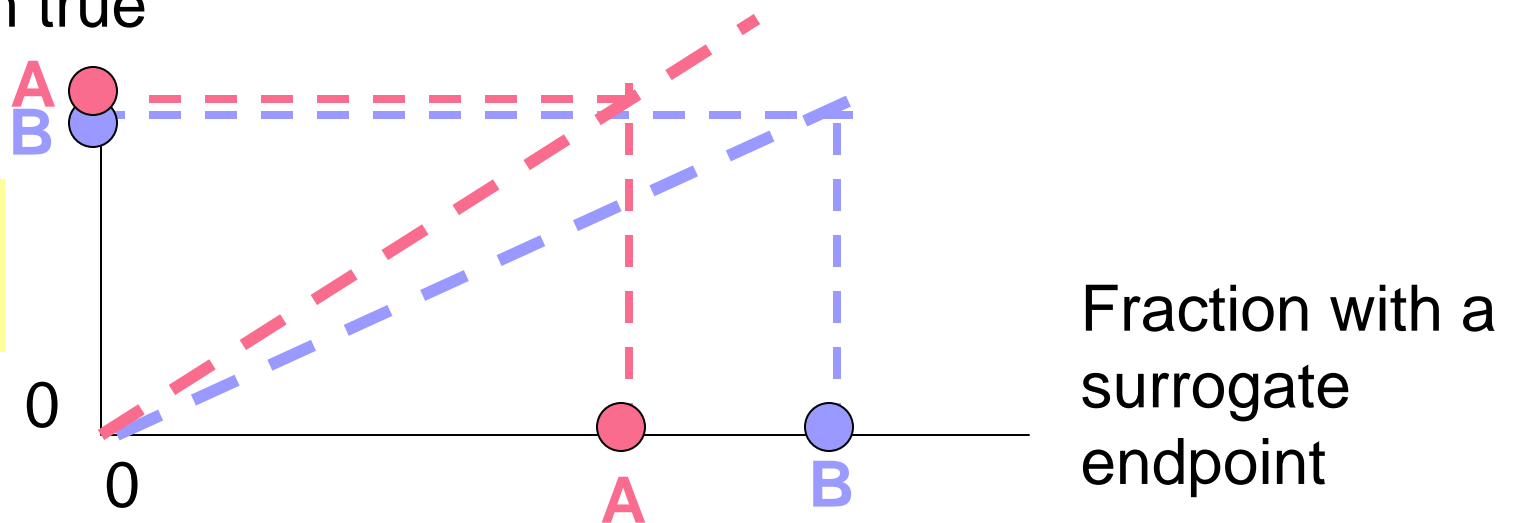
Fraction with true endpoint

Validation trial



Fraction with true endpoint

Application trial



$H_0: A = B$ for true does not imply $H_0: A = B$ for surrogate

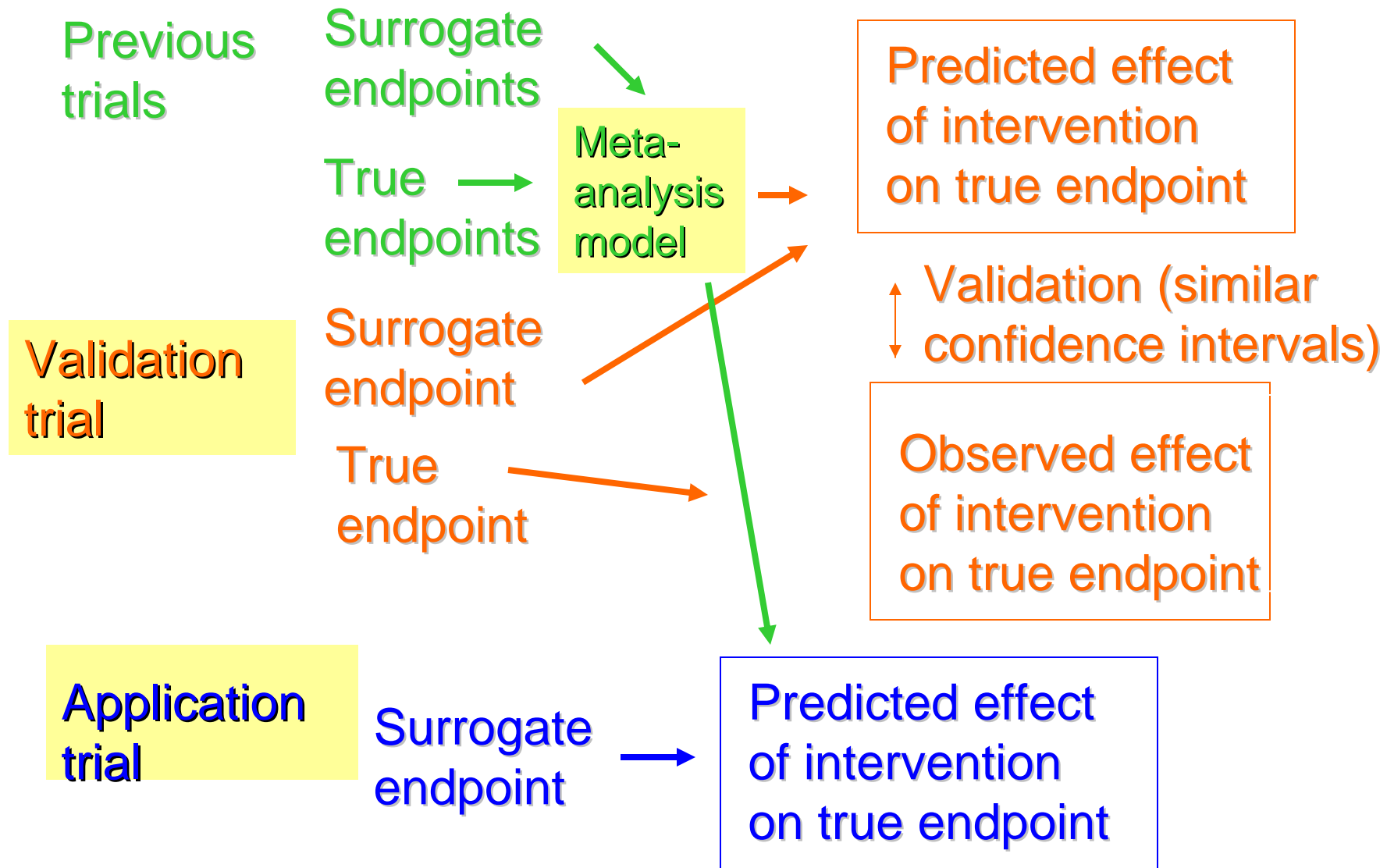
Reject null for surrogate does NOT imply reject null for true



Drawbacks of hypothesis testing framework:

- If Prentice Criterion is not rejected, it is not clear how “close” is good enough.
- With estimation, it is easier to include data from previous trials of surrogate and true endpoint
- Generally estimation is preferred to hypothesis testing (e.g for weighing harms and benefits)

Estimation Framework



DATA SCHEME

T=true; S=surrogate

Application trial

standard treatment

	T=0	T=1
S=0		
S=1		

x
x

new treatment

	T=0	T=1
S=0		
S=1		

x
x

Validation trial

	T=0	T=1
S=0	x	x
S=1	x	x

	T=0	T=1
S=0	x	x
S=1	x	x

Previous trial 1

	T=0	T=1
S=0	x	x
S=1	x	x

	T=0	T=1
S=0	x	x
S=1	x	x

Previous trial 2

	T=0	T=1
S=0	x	x
S=1	x	x

	T=0	T=1
S=0	x	x
S=1	x	x

Previous trial 3

	T=0	T=1
S=0	x	x
S=1	x	x

	T=0	T=1
S=0	x	x
S=1	x	x

“Meta-analytic” methods for combining data from previous trials

- **Method 1** Regression on trial-level statistics
 - e.g. fraction with surrogate and true endpoint in each previous trial
 - Gail et al 2000; Buyse et al 2000
- **Method 2** Combine predicted intervention effects
 - effect of intervention on true endpoint based on data from each previous trial
 - Baker 2005
 - Simpler computations

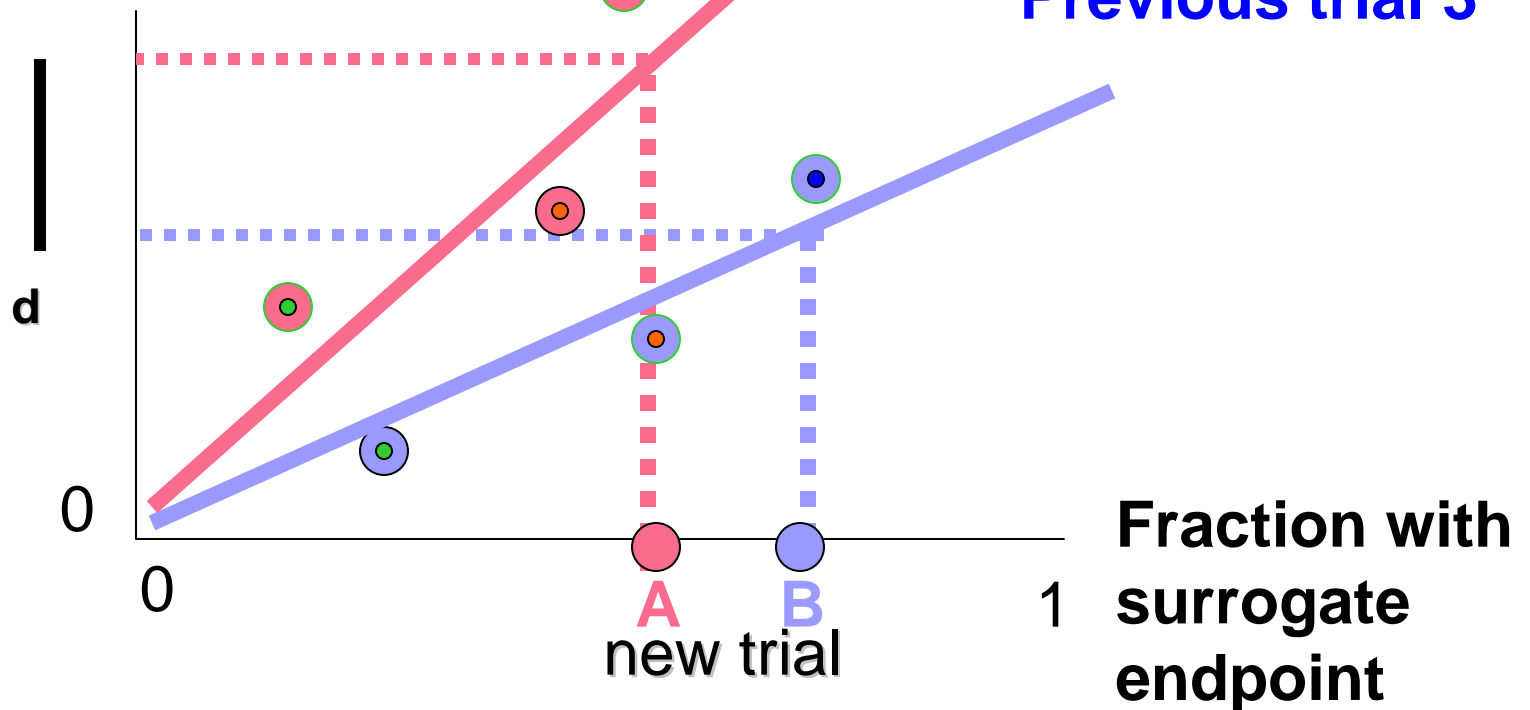
Method 1

“Regression” on trial-level statistics

d is estimated effect of intervention on true endpoint in new trial

Fraction with true endpoint

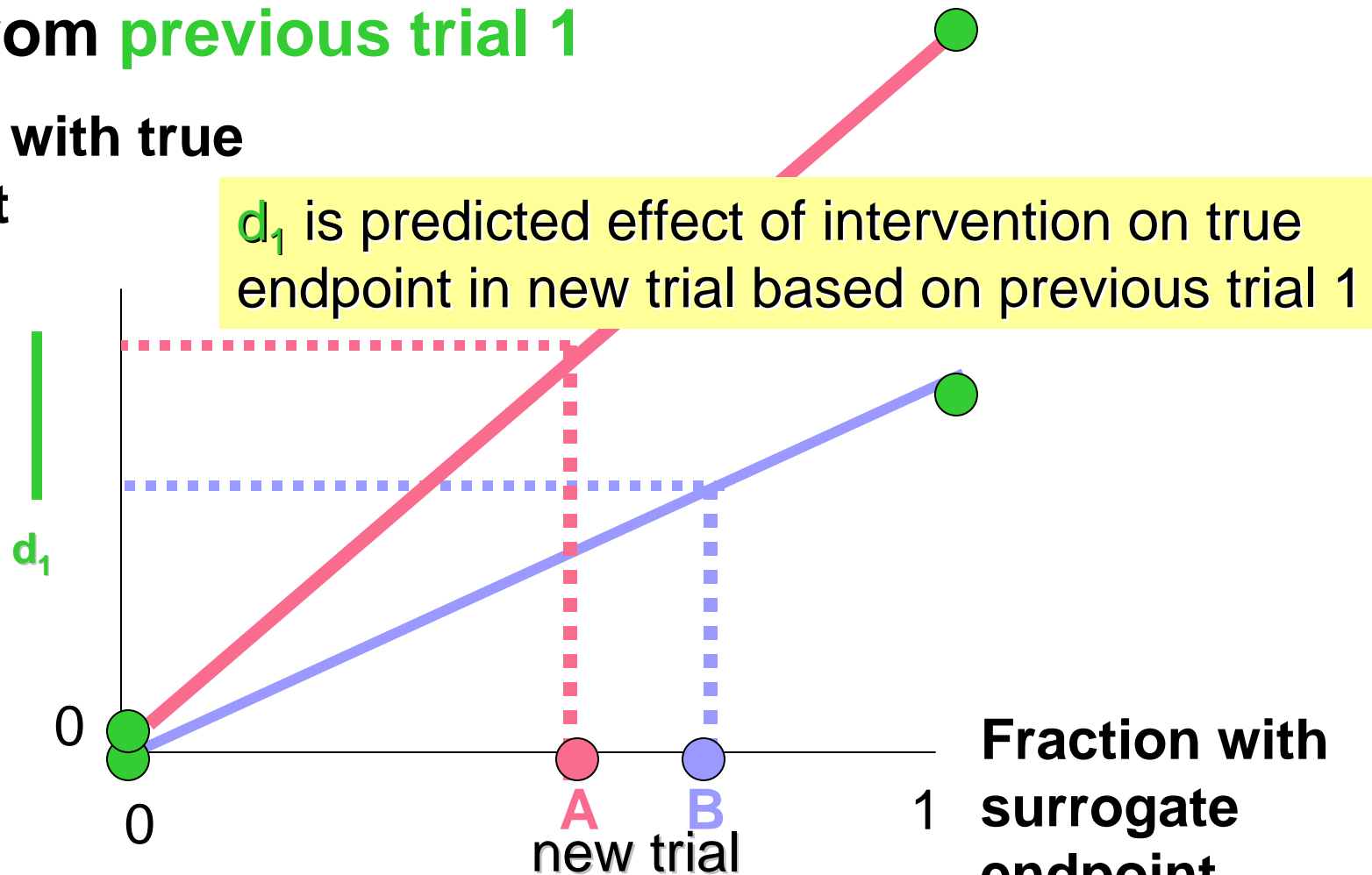
Previous trial 1
Previous trial 2
Previous trial 3



METHOD 2

Predicted effect of intervention on true endpoint based on surrogates **A** and **B** in new trial and data from **previous trial 1**

Fraction with true endpoint



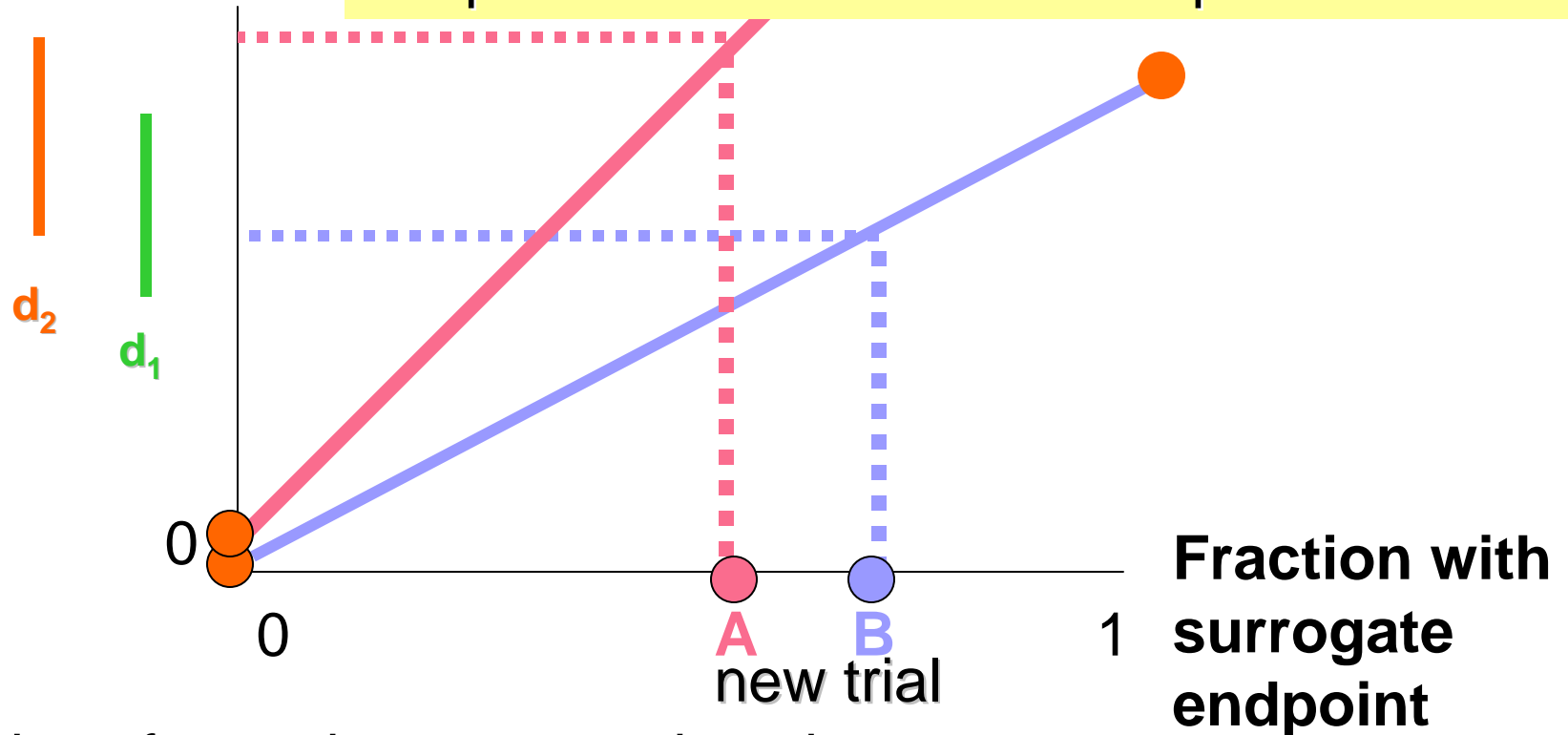
Note: Lines for each group need not be identical — **Prentice Criterion is not needed**

METHOD 2

Predicted effect of intervention on true endpoint based on surrogates **A** and **B** in new trial and data from **previous trial 2**

Fraction with true endpoint

d_2 is predicted effect of intervention on true endpoint in new trial based on previous trial 2

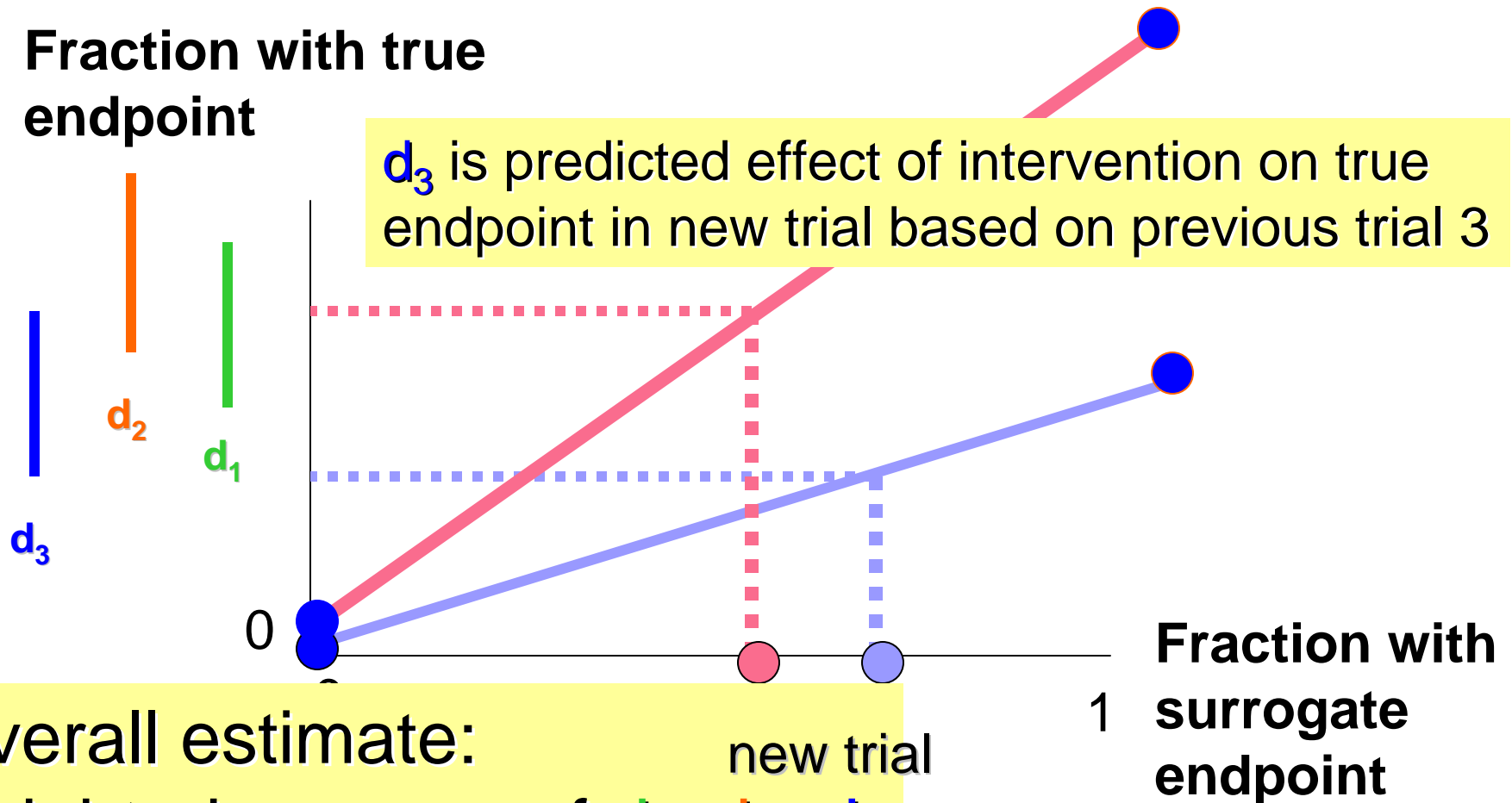


Note: Lines for each group need not be identical — **Prentice Criterion is not needed**

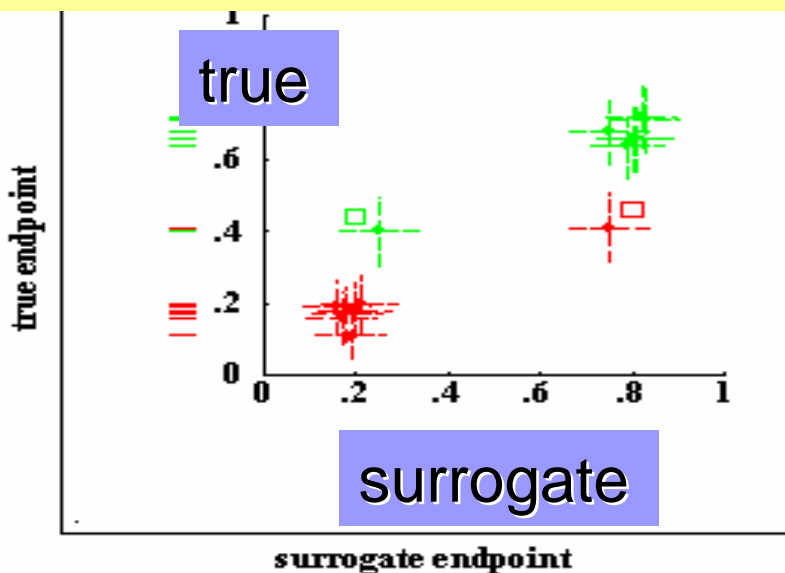
METHOD 2

Predicted effect of intervention on true endpoint based on surrogates **A** and **B** in new trial and data from **previous trial 3**

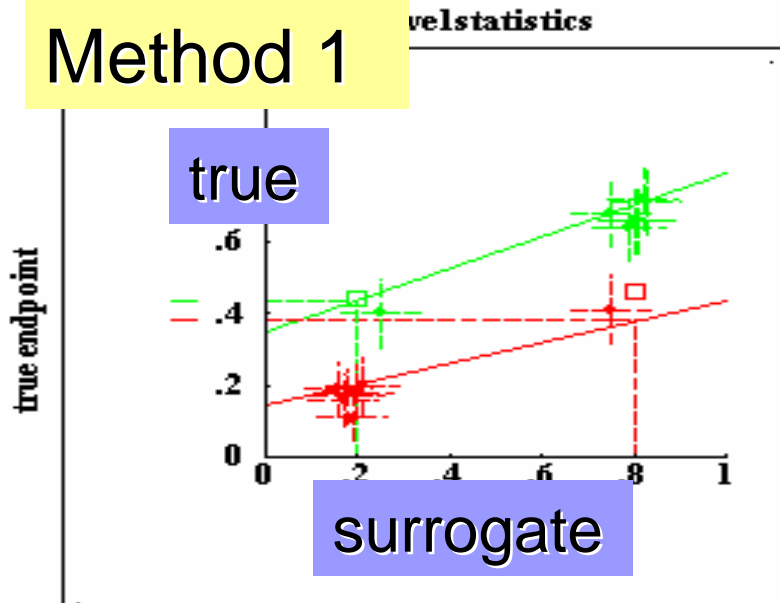
Fraction with true endpoint



HYPOTHETICAL DATA



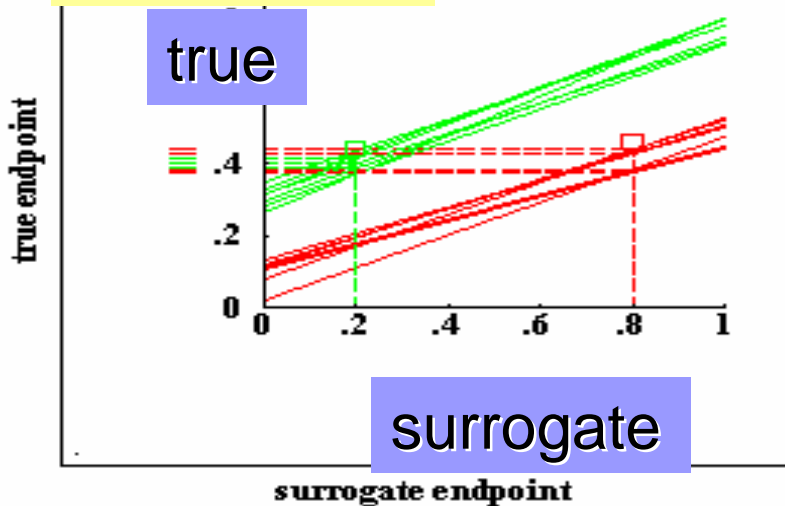
Method 1



Effect of intervention on true endpoint in validation trial

Predicted intervention effects

Method 2

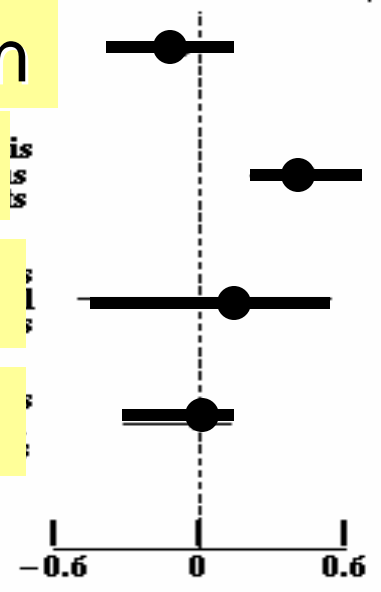


true in validation

previous true

Method 1

Method 2





Be aware of caveats even if surrogate is validated

- Extrapolation to a new intervention
- Surrogate endpoint for benefit does not predict harms that might arise after surrogate is observed



Caveats are less critical if

- Preliminary drug development when next stage is further testing
- Evaluating a different dose or timing of an intervention previously shown effective (using a true endpoint) at another dose or timing



Recommendations

- **Use meta-analytic estimation approach for validation of surrogate endpoint**
 - Check if same conclusion about effect of intervention on true endpoint using (i) surrogate endpoint and (ii) true endpoint
 - But hard to get data from previous trials!
- **Even if validated, remember caveats**
 - extrapolation to a new intervention
 - unknown effect of intervention on harms