Summer Institute in Statistics for Clinical Research

Obtaining insights

to recognize and effectively address scientifically challenging issues in

- Design
- Conduct
- Analysis/Reporting

of clinical trials

Summer Institute in Statistics for Clinical Research

Designs with Active Controls: Non-Inferiority Trials

July 12, 2022

Thomas R. Fleming, Ph.D.
Professor, Dept. of Biostatistics
University of Washington

Reading: Fleming, *Stat in Medicine* 27: 317-332, 2008
Fleming et. al., *Clinical Trials* 8: 432-439, 2011
Fleming & Powers, *J Clin Inf Dis* 47: 108-120, 2008

Pre-Exposure Prophylaxis (PrEP) for HIV: Daily tenofovir/emtricitabine (TDF/FTC) Truvada vs. Placebo

Study	Risk/Gender	Adherence	# of Events	Efficacy; 95% CI
Partners PrEP	Discordant heterosexual couples	~80%	13 vs. 52	75% (55%, 87%)
CDC TDF2	Heterosexual Men/Women	~75%	9 vs. 24	63% (22%, 83%)
iPrEx PROUD iPERGAY	MSM	~60%	41 vs. 97	55% (34%, 69%)
FemPrEP	Heterosexual Women	~35%	33 vs. 35	6% (-69%, 41%)
VOICE	Heterosexual Women	~29%	61 vs. 60	−4% (-50%, 30%)

Alternative strategies in PrEP: Cabotegravir Injectable

- **Longer acting formulation**

  (e.g. *Cabotegravir injectable)*

  – Motivations
    - **Avoid first line treatment drugs**
    - **Lower risk of community resistance**
    - **Somewhat higher** or **similar** efficacy

      through Increased adherence and convenience

  – Safety concerns

Non-Inferiority Trials

- A direct evaluation
  of the clinical efficacy/safety of
  Experimental (**Exp**) relative to *Standard* (**Std**)
  …cannot establish equality…

- Goal: To determine whether
  we can rule out that the efficacy of
  Exp is '*unacceptably worse than*' that of **Std**
  …setting the **Margin**…

E.g.:

- **Cabotegravir (Exp) vs. TDF/FTC (Std)** in PrEP
- **Doripenem (Exp) vs. Piperacillin/Tazo (Std)** in VABP
- **Bivalirudin (Exp) vs. Gp IIb/IIIa (Std)** in PCI

An Important Consideration

- Serious issue if a Standard regimen, established to provide clinically meaningful protection, were to be replaced by a meaningfully less effective intervention

⇒ Reliable evaluation of benefit-to-risk profile of Experimental interventions is necessary…

…this requires development of rigorous evidence-based NI **margins**.

Margin 1.5 ? or 1.125 ?

Dual Goals of Non-Inferiority Trials

- To enable a direct evaluation
  of the clinical efficacy/safety
  of **Exp** relative to **Std**

  …*similarly effective or similarly ineffective?*

- To contribute evidence to
  the evaluation of efficacy/safety
  of **Exp** relative to **Placebo**

E.g.:

- **Cabotegravir (Exp) vs. TDF/FTC (Std)** in PrEP
- **Doripenem (Exp) vs. Piperacillin/Tazo (Std)** in VABP
- **Bivalirudin (Exp) vs. Gp IIb/IIIa (Std)** in PCI

Non-Inferiority Trials… Some Requirements

ICH E9: **Std** should have clinical efficacy

- that is of **substantial magnitude**
- that is **precisely estimated**
- with estimates that are **relevant** to the setting in which the non-inferiority trial is being conducted

Factors invalidating Constancy Assumption
(Exp vs. Std NI Trial vs. Trials evaluating Std)

- ✓ **patient characteristics**

e.g., Disease caused by pathogens resistant to Std in NI Trial

- ✓ **use of supportive care**

e.g., Enhanced concomitant Rx attenuates effect of Std in NI Trial

- ✓ dose, schedule, level of adherence

e.g., Lower adherence to Std in NI trial

- ✓ efficacy and safety endpoints

~ definition ~ validation process ~ missing data

.....as in maintaining conditions of a lab experiment…

Factors invalidating Constancy Assumption

✓ *use of supportive care*

7/16/08 Anti-Infective Drugs Advisory Committee

DORI - 09

Dori — Adjunctive pseudomonal Rx: ≈ 80%

Pip / Tazo — Adjunctive anti-MRSA Rx: ≈ 15%

45% of Dori pts received i.v. & oral therapy

…..FDA: *"The evaluation of clinical response for most patients is confounded by the prolonged use of adjunctive amikacin therapy"*

...among 109 clinically evaluable cures on Doripenem,

≥ 39 rec'd single agent Doripenem ≤ 2 days…

…FDA: *"discuss how the treatment effect of study drug will be determined in patients administered combination antibacterial therapy"*

Factors invalidating Constancy Assumption

✓ *use of supportive care*

Daptomycin vs. Ceftriaxone in CABP

Clinical Cure Rate in Clinically Evaluable Population

Pertel et al
CID 46: 1142-1151, 2008

	Overall		Prior Effective Antibacterial Therapy: Yes		Prior Effective Antibacterial Therapy: No	
	n	C.R.	n	C.R.	n	C.R.
✓ Daptomycin	369	79.4%	97	90.7%	272	75.4%
✓ Ceftriaxone	371	87.9%	92	88.0%	279	87.8%
(95% C.I.)	*(**-13.8**, -3.2)*		*(**-6.1**, 11.5)*		*(-18.8, **-6.0**)*	

"Daptomycin is not effective for the Rx of CABP…trials to evaluate CABP Rx may need to exclude patients who have rec'd any potentially effective prior Rx"

Factors invalidating Constancy Assumption
*(**Exp** vs. **Std** NI Trial vs. Trials evaluating **Std**)*

- ✓ patient characteristics

e.g., Participants less likely to be impacted by **Std** in NI Trial

- ✓ use of supportive care

e.g., Enhanced concomitant Rx attenuates effect of **Std** in NI Trial

- ✓ **dose, schedule, level of adherence**

e.g., Lower adherence to **Std** in NI trial

- ✓ efficacy and safety endpoints

~ well-defined & reliable *~ clinically meaningful* *~ sensitive*

Pre-Exposure Prophylaxis (PrEP): Daily tenofovir/emtricitabine (TDF/FTC) Truvada vs. Placebo

Study	Risk/Gender	Adherence	# of Events	Efficacy; 95% CI
Partners PrEP	Discordant heterosexual couples	~80%	13 vs. 52	75% (55%, 87%)
CDC TDF2	**Heterosexual Men/Women**	~75%	9 vs. 24	63% (22%, 83%)
iPrEx PROUD iPERGAY	MSM	~60%	41 vs. 97	55% (34%, 69%)
FemPrEP	**Heterosexual Women**	~35%	33 vs. 35	6% (-69%, 41%)
VOICE	Heterosexual Women	~29%	61 vs. 60	−4% (-50%, 30%)

Factors invalidating Constancy Assumption
*(**Exp** vs. **Std** NI Trial vs. Trials evaluating **Std**)*

- ✓ patient characteristics

e.g., Participants less likely to be impacted by **Std** in NI Trial

- ✓ use of supportive care

e.g., Enhanced concomitant Rx attenuates effect of **Std** in NI Trial

- ✓ dose, schedule, level of adherence

e.g., Lower adherence to **Std** in NI trial

- ✓ **efficacy and safety endpoints**

~ well-defined & reliable *~ clinically meaningful* *~ sensitive*

How to Achieve Scientific Objectivity in Selecting Trials to Estimate Efficacy of Std?

Illustration: Pemetrexed (Exp) *vs* Docetaxel (Std)
in 2nd Line NSCLC patients
(Overall Survival)

Non Inferiority Trial	**Death**	**Median Survival**
Alimta (Pemetrexed)	206/283	8.3 mo
Docetaxel (75 mg/m^2)	203/288	7.9 mo

RR = 0.992 (0.82, 1.20)

Two Trials	**TAX 317**		**TAX 320**	
	N Surv	N Surv	Deaths	Med Surv
Docetaxel 100 (mg/m^2)	49 6.0 m	—	97/ 125	5.7 m
Docetaxel 75 (mg/m^2)	—	55 8.0 m	104/ 125	5.5 m
Best Supportive Care*	51 5.0 m	49 4.7 m	110/ 123	5.6 m
	RR ≈ 0.95	**RR = 0.56**		

***analgesics, radiotherapy**

*** vinorelbine or ifosfamide**

In choosing evidence to estimate the efficacy of Std

A process is needed

that will provide greater assurance of Scientific Objectivity in the determination of:

~ The proper historical studies

~ The proper sub-samples from these studies

Illustration: Setting the Margin

Injectable (**Exp**) *vs* TDF/FTC (**Std**)
PrEP in MSM
(Rate of HIV Infection)

NI Trial (e.g. HPTN 083) **HIV INFECTION**

Injectable
TDF/FTC

Factors Influencing the Choice of Margin and Interpretation of NI Trial Results

- Active Control (i.e., Std) Effect
- Clinical Relevance of:

  Loss of *Benefit* (i.e. **3 add'l HIV inf / 1000 p.y.**)

  relative to changes in:

  Fewer side effects

  Avoid first line treatment drugs

  Lower risk of community resistance

Non-Inferiority Trials… Some Requirements

ICH E9: **Std** should have clinical efficacy

- that is of **substantial magnitude**
- that is **precisely estimated**
- with estimates that are **relevant** to the setting in which the non-inferiority trial is being conducted

Pre-Exposure Prophylaxis (PrEP): Daily tenofovir/emtricitabine (TDF/FTC) Truvada vs. Placebo

Study	Risk/Gender	Adherence	# of Events	Efficacy; 95% CI
Partners PrEP	Discordant heterosexual couples	~80%	13 vs. 52	75% (55%, 87%)
CDC TDF2	Heterosexual Men/Women	~75%	9 vs. 24	63% (22%, 83%)
iPrEx PROUD iPERGAY	**MSM**	~60%	**41 vs. 97**	55% (34%, 69%)
FemPrEP	Heterosexual Women	~35%	33 vs. 35	6% (-69%, 41%)
VOICE	Heterosexual Women	~29%	61 vs. 60	−4% (-50%, 30%)

Illustration: Setting the Margin

Injectable (**Exp**) *vs* TDF/FTC (**Std**)
PrEP in MSM
(Rate of HIV Infection)

NI Trial (e.g. HPTN 083) **HIV INFECTION**

Injectable
TDF/FTC

iPrEx/PROUD/iPERGAY **HIV INFECTION**

TDF/FTC	Total events
Placebo	≈ 138

(TDF/FTC / Placebo) RR = 0.45 95% CI: (0.31, 0.66)

"HIV Infection" Events

Placebo compared with TDF/FTC

Placebo better

TDF/FTC better

Illustration: Setting the Margin

Injectable (**Exp**) *vs* TDF/FTC (**Std**)
PrEP in MSM
(Rate of HIV Infection)

NI Trial (e.g. HPTN 083) **HIV INFECTION**

Injectable
TDF/FTC

iPrEx/PROUD/iPERGAY **HIV INFECTION**

TDF/FTC	Total events
Placebo	≈ 138

(TDF/FTC / Placebo) RR = 0.45 95% CI: (0.31, 0.66)

Illustration: Setting the Margin

Injectable (**Exp**) *vs* TDF/FTC (**Std**)
PrEP in MSM
(Rate of HIV Infection)

NI Trial (e.g. HPTN 083) **HIV INFECTION**

Injectable
TDF/FTC

iPrEx/PROUD/iPERGAY **HIV INFECTION**

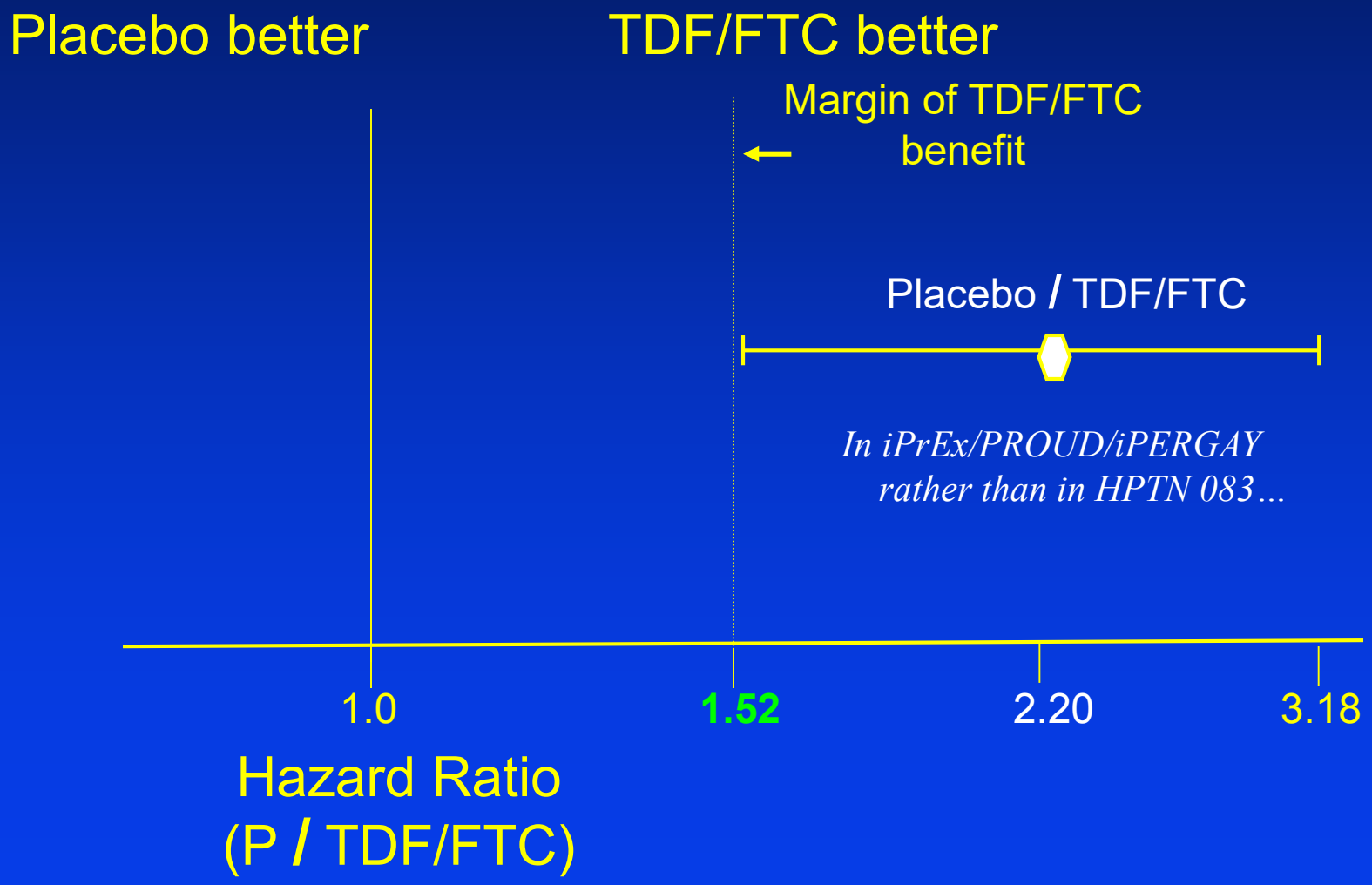
TDF/FTC	Total events
Placebo	≈ 138

(TDF/FTC / Placebo) RR = 0.45 95% CI: (0.31, 0.66)

(Placebo / TDF/FTC) RR = 2.20 95% CI: (**1.52**, 3.18)

"HIV Infection" Events

Placebo compared with TDF/FTC



Factors invalidating Constancy Assumption
(Non-Inferiority Trial vs. iPrEx/PROUD/iPERGAY)

- ✓ patient characteristics

e.g., Participants less likely to be impacted by **Std** in NI Trial

- ✓ use of supportive care

e.g., Enhanced concomitant Rx attenuates effect of **Std** in NI Trial

- ✓ dose, schedule, level of adherence

e.g., Lower adherence to **Std** in NI trial

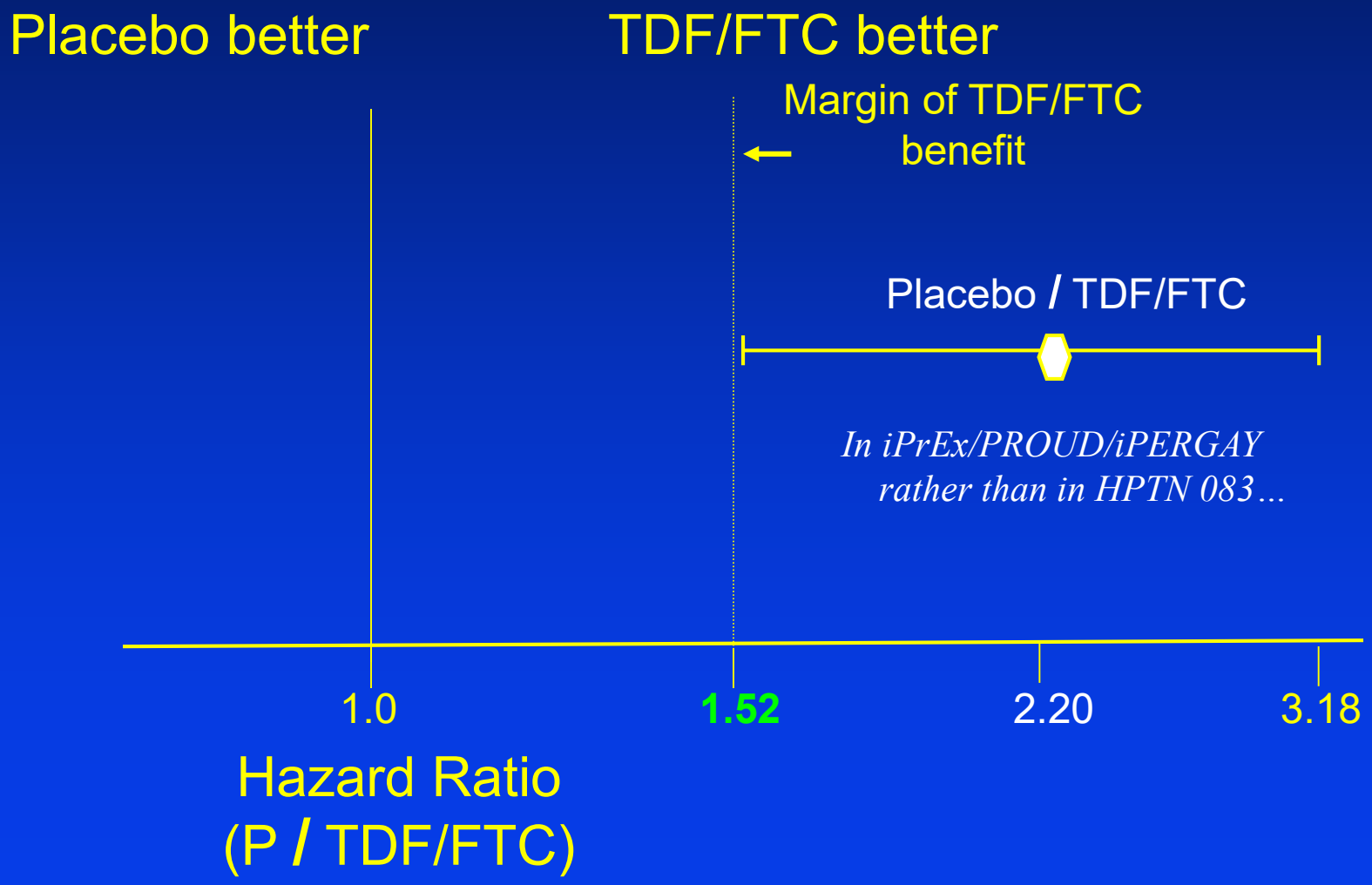
- ✓ efficacy and safety endpoints

~ definition ~ validation process ~ missing data

.....as in maintaining conditions of a lab experiment…

"HIV Infection" Events

Placebo compared with TDF/FTC



Factors Influencing Choice of Margin

- Active Control (i.e. **Std**) Effect

  (on risk of HIV Infection)

~ magnitude of Active Control effect

Eg: Estimated (P / TDF/FTC) Relative Risk = 2.20

~ precision of estimate

Eg: ± 2 s.e. = (**1.52**, 3.18) (138 events)

~ estimates relevant to setting of NI trial

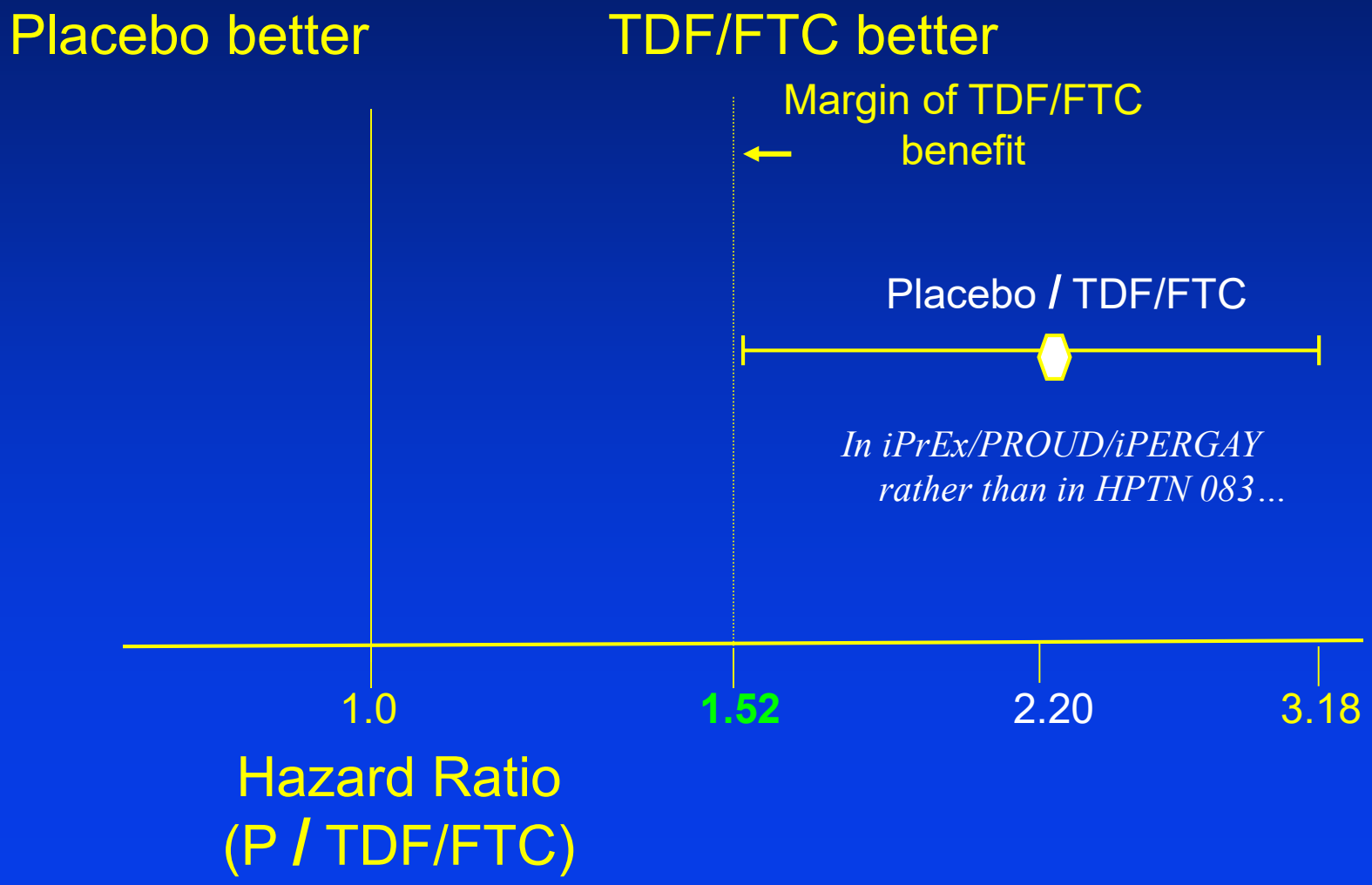
- Population
- Adherence
- Supportive care
- Endpoint assessment

~ preserve > half of the Active Control effect

$\sqrt{\mathbf{1.52}} = \mathbf{1.23}$

"HIV Infection" Events

Placebo compared with TDF/FTC



Factors Influencing Choice of Margin

- Active Control (i.e. **Std**) Effect

  (on risk of HIV Infection)

~ magnitude of Active Control effect

Eg: Estimated (P / TDF/FTC) Relative Risk = 2.20

~ precision of estimate

Eg: ± 2 s.e. = (**1.52**, 3.18) (138 events)

~ estimates relevant to setting of NI trial

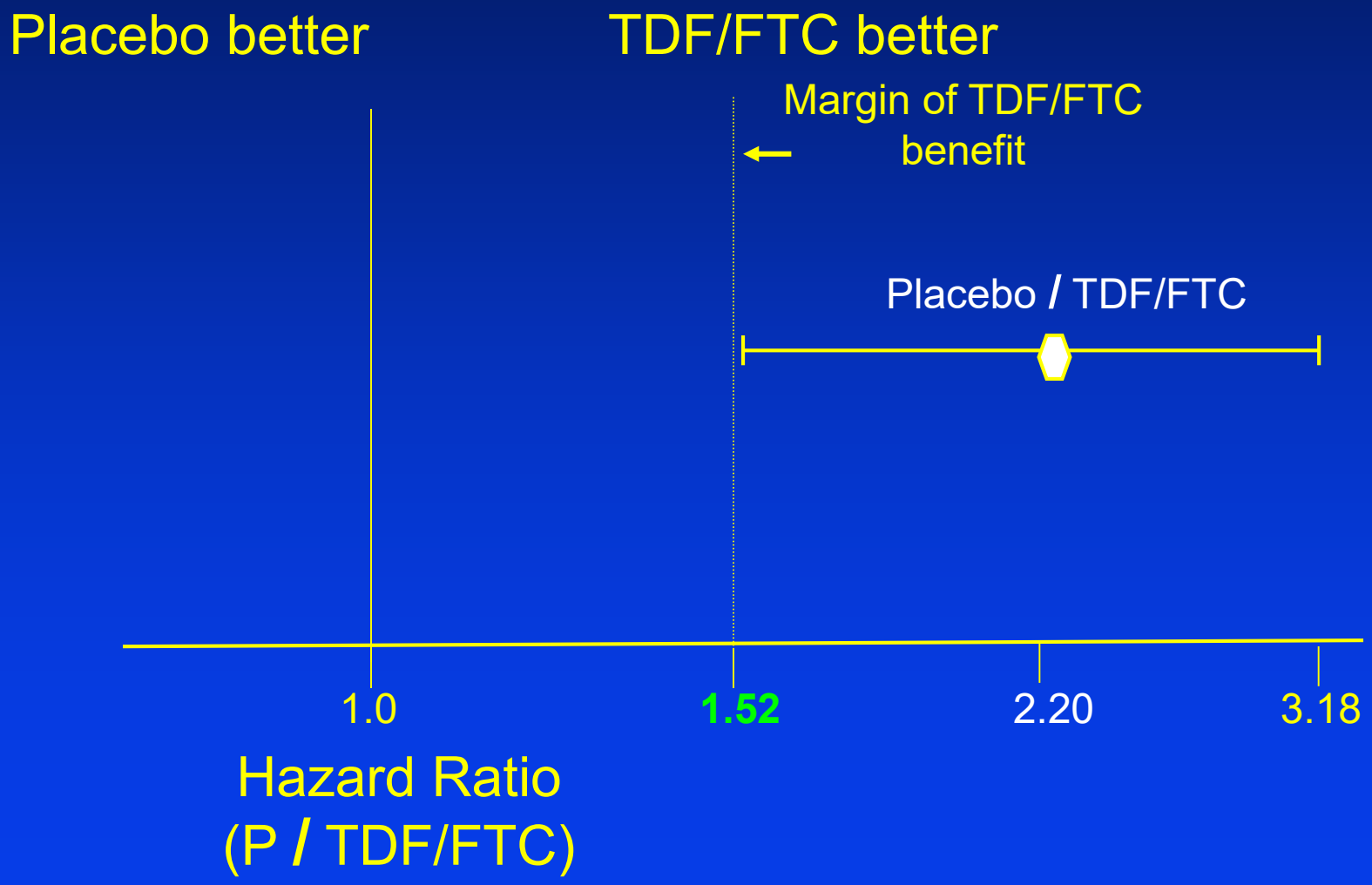
- Population
- Adherence
- Supportive care
- Endpoint assessment

~ preserve > half of the Active Control effect

$\sqrt{\mathbf{1.52}} = \mathbf{1.23}$

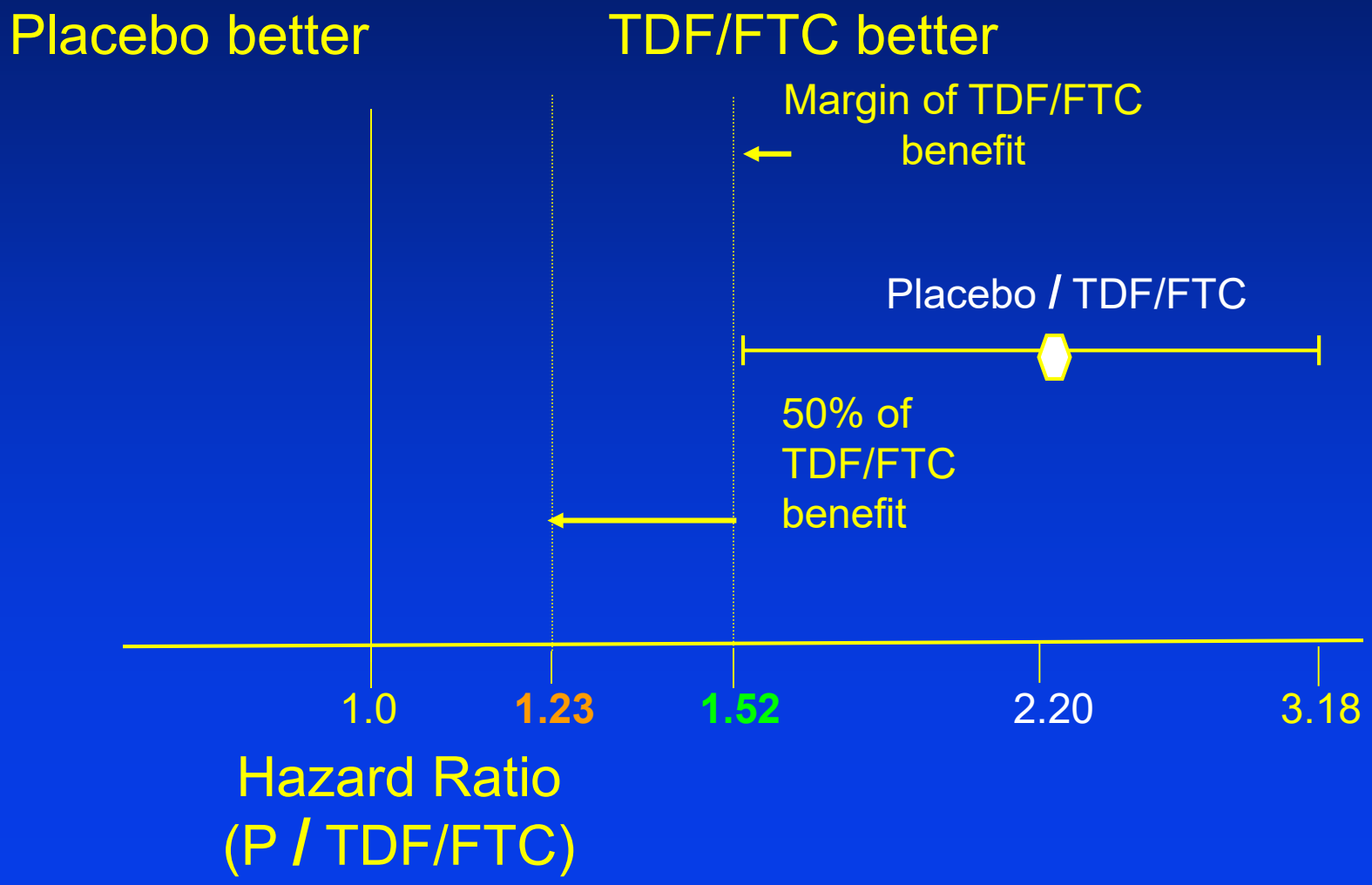
"HIV Infection" Events

Placebo compared with TDF/FTC



"HIV Infection" Events

Placebo compared with TDF/FTC



Factors Influencing the Choice of Margin and Interpretation of NI Trial Results

- Active Control (i.e., Std) Effect
- Clinical Relevance of:

  Loss of *Benefit* (i.e. 3 add'l HIV inf / 1000 p.y.)

  relative to changes in:

  Fewer side effects
  Avoid first line treatment drugs
  Lower risk of community resistance

Illustration: Setting the Margin

Injectable (**Exp**) *vs* TDF/FTC (**Std**)
PrEP in MSM
(Rate of HIV Infection)

NI Trial (e.g. HPTN 083) **HIV INFECTION**

Injectable
TDF/FTC

1.25/100 p.y.

iPrEx Trial **HIV INFECTION**

TDF/FTC
Placebo

Total events
≈ 138

(TDF/FTC / Placebo) RR = 0.45 95% CI: (0.31, 0.66)

(Placebo / TDF/FTC) RR = 2.20 95% CI: (**1.52**, 3.18)

Factors Influencing the Choice of Margin and Interpretation of NI Trial Results

- Active Control (i.e., Std) Effect
- Clinical Relevance of:

  Loss of *Benefit* (i.e. **3 add'l HIV inf / 1000 p.y.**)

  relative to changes in:

  Fewer side effects
  Avoid first line treatment drugs
  Lower risk of community resistance

"HIV Infection" Events

Placebo compared with TDF/FTC

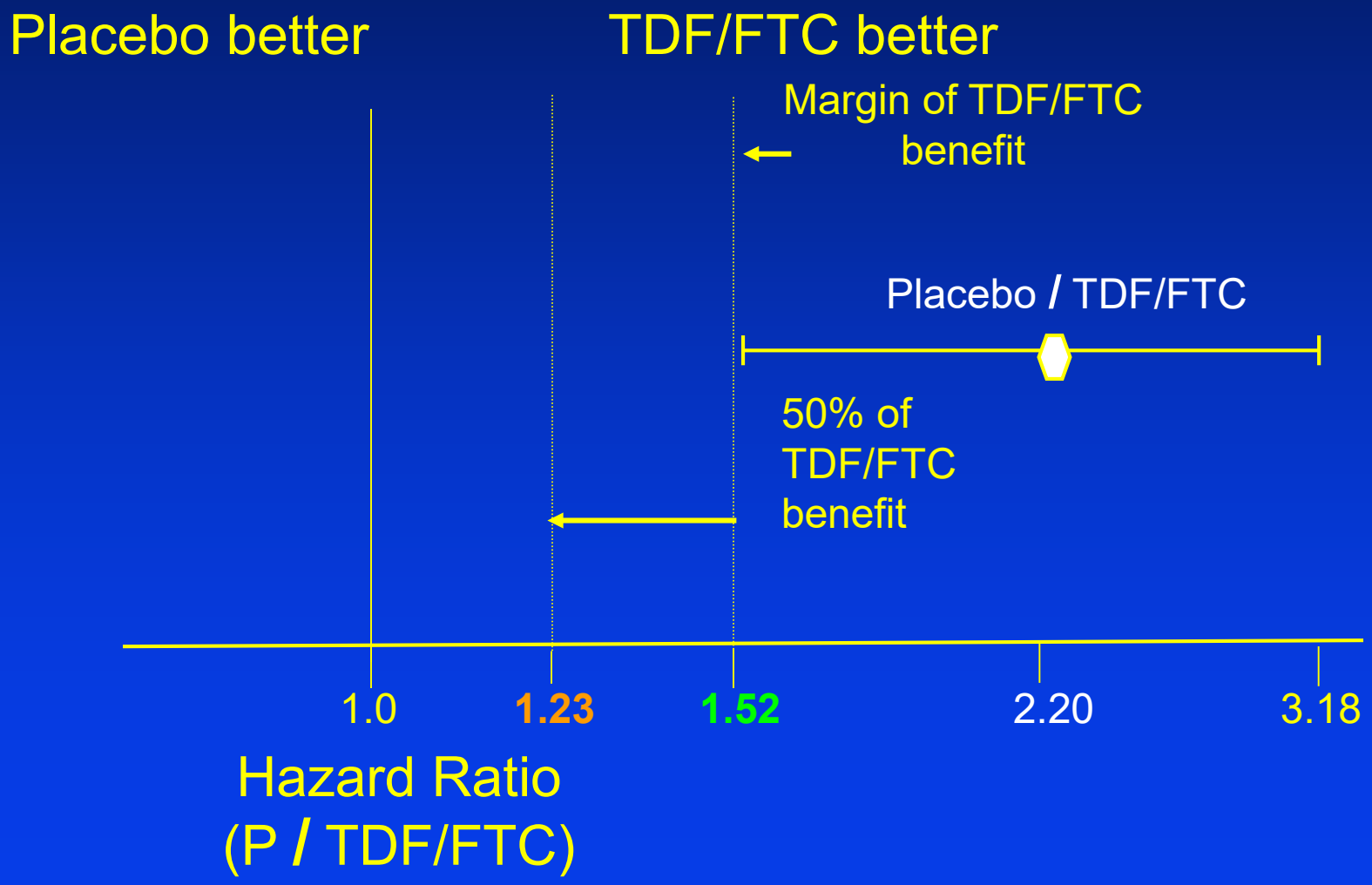


Illustration: Setting the Margin

Injectable (**Exp**) *vs* TDF/FTC (**Std**)
PrEP in MSM
(Rate of HIV Infection)

NI Trial (e.g. HPTN 083) **HIV INFECTION**

Injectable
TDF/FTC

1.25/100 p.y.

iPrEx Trial **HIV INFECTION**

TDF/FTC
Placebo

Total events
≈ 138

(TDF/FTC / Placebo) RR = 0.45 95% CI: (0.31, 0.66)

(Placebo / TDF/FTC) RR = 2.20 95% CI: (**1.52**, 3.18)

Illustration: Setting the Margin

Injectable (**Exp**) *vs* TDF/FTC (**Std**)
PrEP in MSM
(Rate of HIV Infection)

NI Trial (e.g. HPTN 083) **HIV INFECTION**

Injectable	**13** / 3171 p.y. f.u.	
TDF/FTC	**39** / 3197 p.y. f.u.	
	RR = 0.34 (0.16, **0.62**)	1.25/100 p.y.

iPrEx Trial **HIV INFECTION**

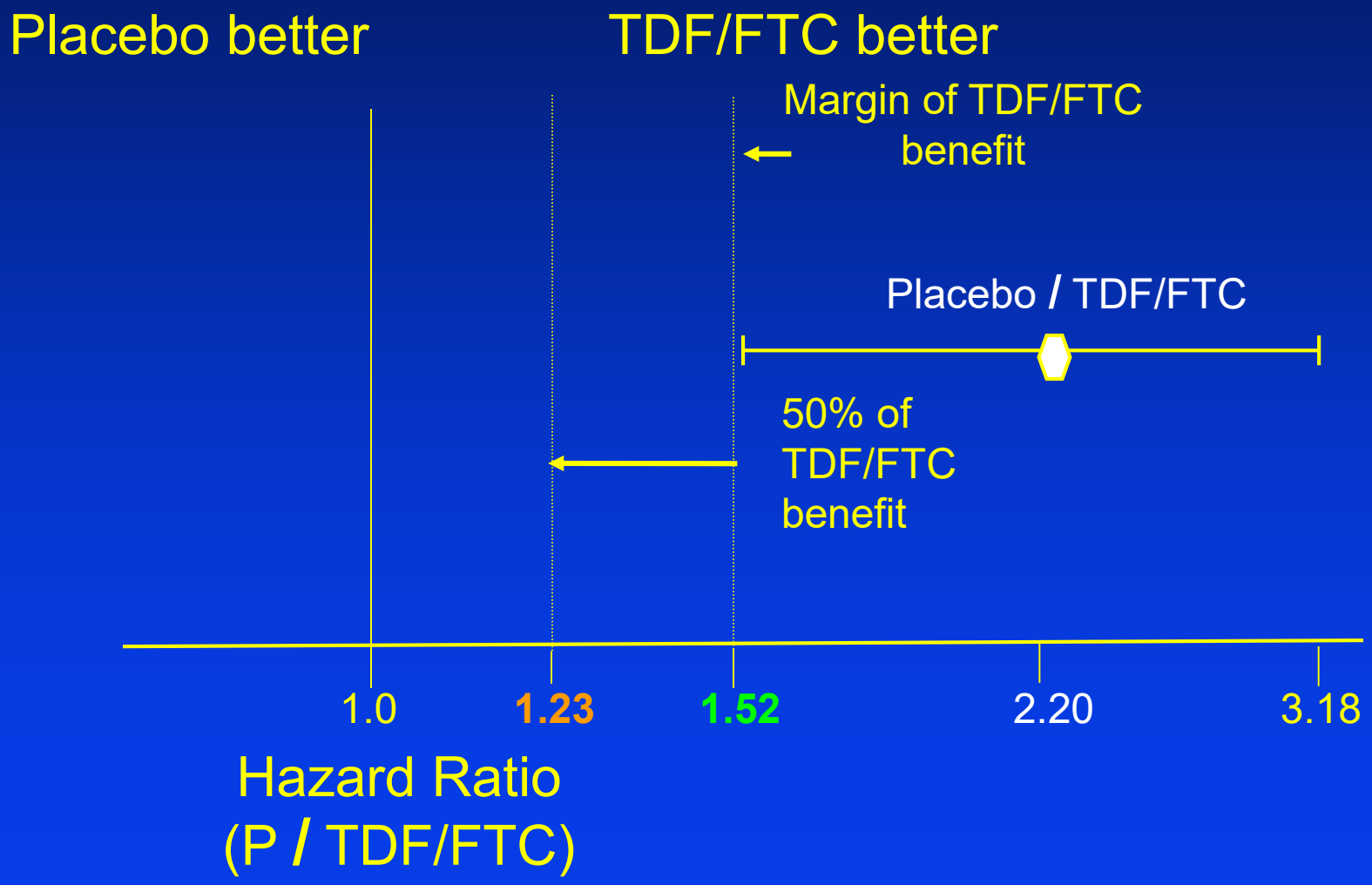
TDF/FTC	Total events
Placebo	≈ 138

(TDF/FTC / Placebo) RR = 0.45 95% CI: (0.31, 0.66)

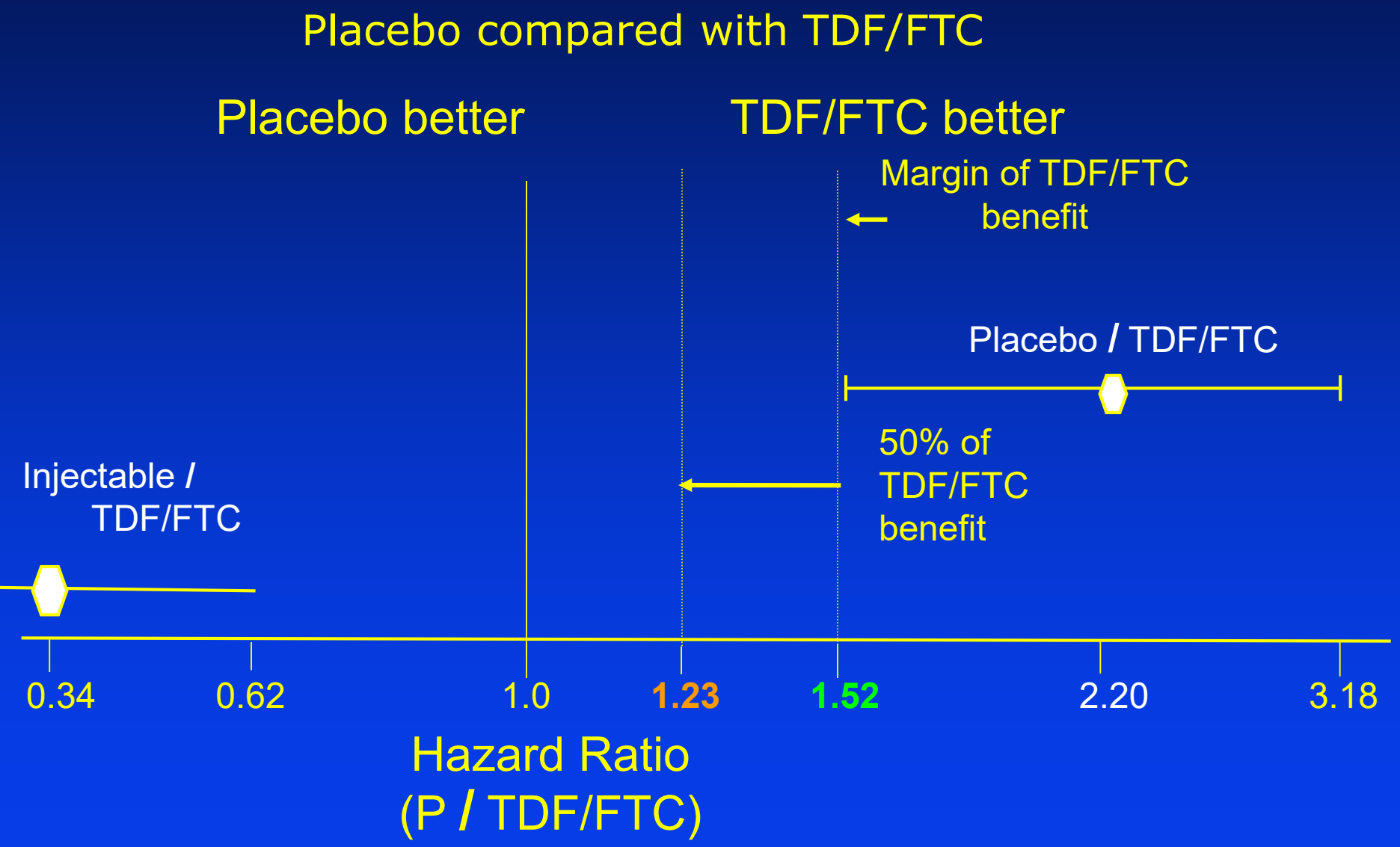
(Placebo / TDF/FTC) RR = 2.20 95% CI: (**1.52**, 3.18)

"HIV Infection" Events

Placebo compared with TDF/FTC



"HIV Infection" Events



Determining the Margin in NI Trials

Goal in NI trials: Ruling out the new intervention (**Exp**) is unacceptably worse than a standard (**Std**) regimen having *reliable* evidence of *substantial* effects…

⇒ Need an 'evidence based' NI **Margin**

Determining the NI margin: Two Key considerations

- The NI margin should be formulated using adjustments to account for bias or inherent unreliability in the estimate of the effect of **Std** in the non-inferiority trial setting.
  (…*as in superiority trials that are not randomized*…)
- The NI margin should be formulated to preserve an appropriate percentage of the effect of **Std**.

Community Acquired Pneumonia: Mortality
(Non-bacteremic patients, Age > 50)

- *Sulfonamide derivatives & penicillin. (Fleming, Powers. *CID*, 2008)

	21-day Mortality
➢ Antibiotics*	16.1%
➢ No Specific Rx	49.4%

- Consider an **Exp** *in patients who are candidates for Antibiotics*:

	21-day Mortality
➢ Experimental Rx	37%
➢ No Specific Rx	49%

- Is a statistically significant, but clinically modest, ↓ in mortality acceptable *in patients who are candidates for Antibiotics*?

Clinton-Gore (April 1995)

– "it is essential for public health protection that a new therapy be as effective as alternatives that are already approved for marketing when:

1. the disease to be treated is life-threatening or capable of causing irreversible morbidity (e.g., stroke or heart attack); or
2. the disease to be treated is a contagious illness that poses serious consequences to the health of others (e.g., sexually transmitted disease)."

The Choice of the Margin in a NI Trial

ICH E10: "The determination of the **margin** in a non-inferiority trial is based on both *statistical reasoning* & *clinical judgment*, and should reflect uncertainties in the evidence on which the choice is based, and should be *suitably conservative*."

The Utility of NI Trials in Clinical Research

"Non-inferiority trials with non-rigorous margins allow substantial risk for accepting inadequately effective experimental regimens, leading to the risk of erosion in quality of health care…

Due to the inherent uncertainties in non-inferiority trials, alternative designs should be pursued whenever possible."

* Fleming TR, Odem-Davis K, Rothmann MD, Shen YL "Some essential considerations in the design and conduct of non-inferiority trials." *Clinical Trials* 8: 432-439, 2011

NI margin 1.8 $\Rightarrow$ need est HR $\leq$ 1.31 NI margin 1.3 $\Rightarrow$ need est HR $\leq$ 1.10

Non-Inferiority Trials

Summary
and
Recommendations

Non-Inferiority Trials

- Do not establish Exp is "as effective as" Std;
  …NI trials rule out Exp is "unacceptably worse"
- Margins should be smaller than
  differences in efficacy that patients & caregivers
  consider to be clinically relevant
- Margins should not be based on what can be
  ruled out using a pre-specified sample size
  (1993 FDA Anti-Infective Drugs Guidance Document)

Doripenem vs. Piperacillin/Tazo…15% margin?
Cabotegravir vs. Truvada…1.5 or 1.125?

Non-Inferiority Trials

- Bio-creep can be avoided without
  necessarily requiring huge sample sizes
- NI Trials with Surrogate Endpoints:
  Treacherous!
- NI trial designs should be avoided if possible…
  …they share many of the inherent dangers
  of historically controlled trials….

Garattine S, Bertele V. "NI trials are unethical because they disregard patients' interests." *Lancet* 2007; 370: 1875-77

Non-Inferiority Trials

- Best motivation when experimental regimen
  has favorable profile in
  side effects, cost,
  or convenience of administration
- Standard (Std) should have clinical efficacy
  that is
  - of substantial magnitude
  - precisely estimated in the population
    from which the study sample is drawn
- This design approach imposes constraints that
  the NI trial be conducted in a setting similar to that
  of the trials used to estimate the effect of Std

Non-Inferiority Trials vs. Superiority Trials

- ICH E10: "The determination of the margin in a non-inferiority trial is based on both *statistical reasoning* & *clinical judgment*, should reflect uncertainties in the evidence on which the choice is based, and should be *suitably conservative*."

- When one cannot justify a non-trivial margin, randomized controlled superiority trials provide an ethically and scientifically reliable approach to assessing the benefit-to-risk profile

The Utility of NI Trials in Clinical Research

"Non-inferiority trials with non-rigorous margins allow substantial risk for accepting inadequately effective experimental regimens, leading to the risk of erosion in quality of health care…

Due to the inherent uncertainties in non-inferiority trials, alternative designs should be pursued whenever possible."

* Fleming TR, Odem-Davis K, Rothmann MD, Shen YL "Some essential considerations in the design and conduct of non-inferiority trials." *Clinical Trials* 8: 432-439, 2011