Sampling for an Effectiveness Study
OR
“How to reject your most hated hypothesis”

Mead Over and Sergio Bautista

MC Evaluation Workshop and Operations Meeting
January 18-23, 2010
Johannesburg, South Africa

Hypothesis to reject

- Circumcision has no impact
- Circumcision has too little impact
- Intensive Circumcision Program has no more impact than Routine Circumcision Program
- Circumcision has no benefit for women

Hypothesis to reject

- **Circumcision has no impact**
- Circumcision has too little impact
- Intensive Circumcision Program has no more impact than Routine Circumcision Program
- Circumcision has no benefit for women

Efficacy

Problem, we might get it wrong: Power

Ho = MC does not impact HIV incidence

		State of the World	
		MC does not change HIV incidence	MC changes HIV incidence
Estimate	MC does not change HIV incidence	Correct acceptance of Ho	**Power: probability that you don't fail to reject no impact, that you find impact** Incorrect acceptance of Ho (wrongly saying MC didn't work)
	MC reduces HIV incidence	Incorrect rejection of Ho (wrongly saying MC did work)	Correct rejection of Ho

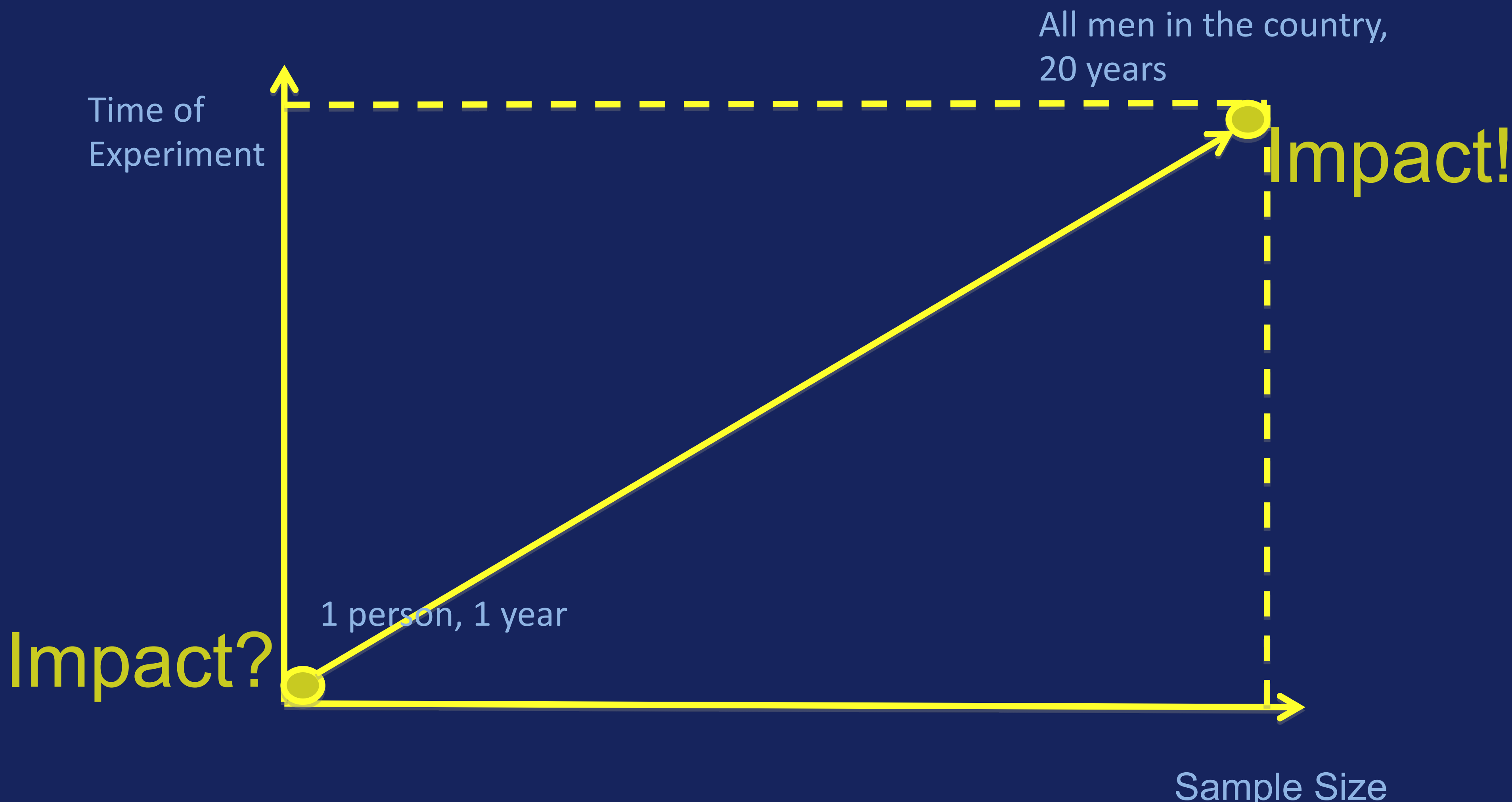
Confidence, power, and two types of mistakes

- Confidence describes the test's ability to minimize type-I errors (false positives)
- Power describes the test's ability to minimize type-II errors (false negatives)
- Convention is to be more concerned with type-I than type-II errors
  - (ie, more willing to mistakenly say that something didn't work when it actually did, than to say that something worked when it actually didn't)
- We usually want confidence to be 90 – 95%, but will settle for power of 80 – 90%

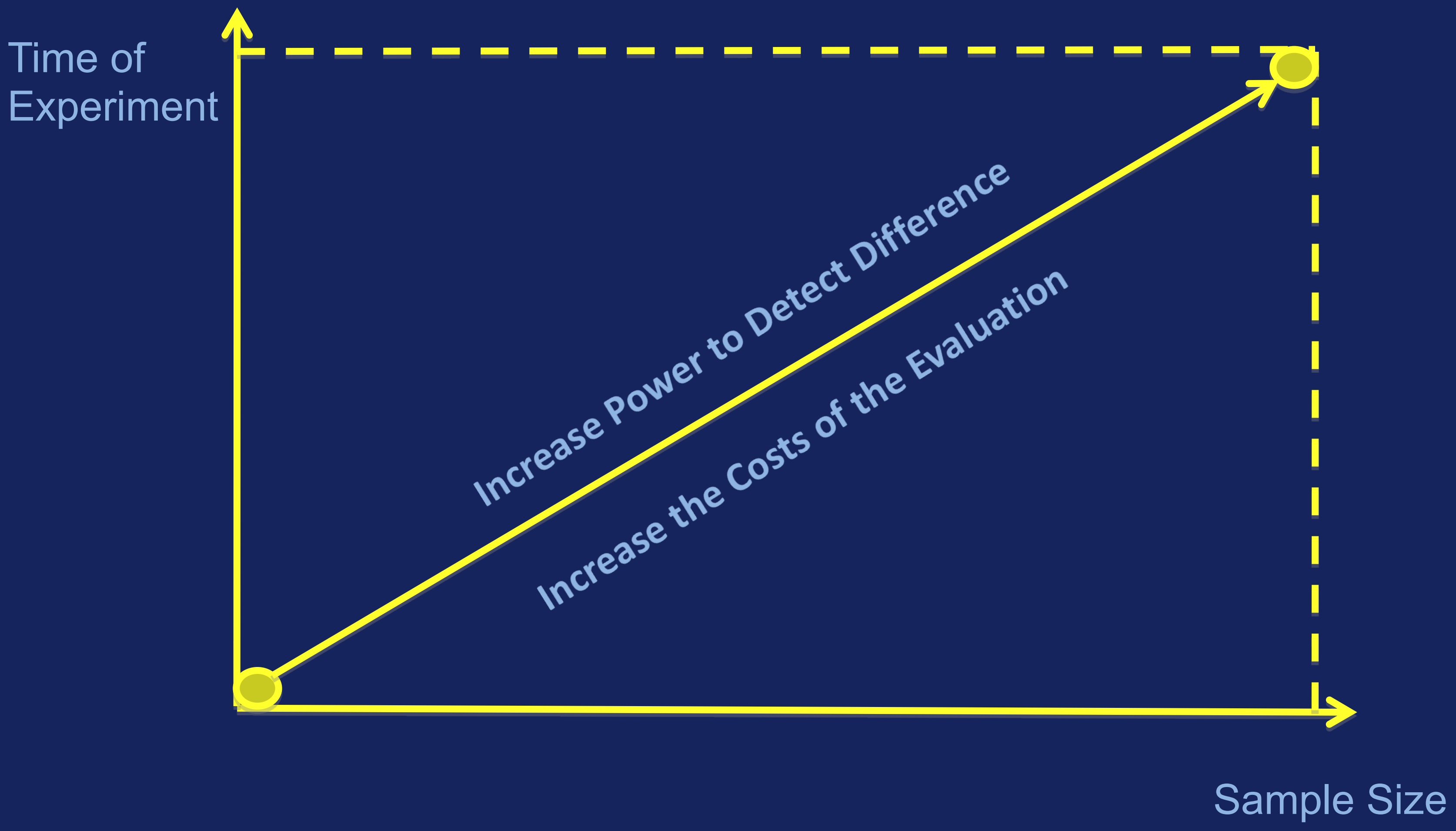
Power

- As power increases, the chances to say “no impact” when in reality there is positive impact, declines
- Power analysis can be used to calculate the minimum sample size required to accept the outcome of a statistical test with a particular level of confidence

The problem


All men in the country,
20 years
Time of
Experiment
Impact!
1 person, 1 year
Impact?
Sample Size

The problem

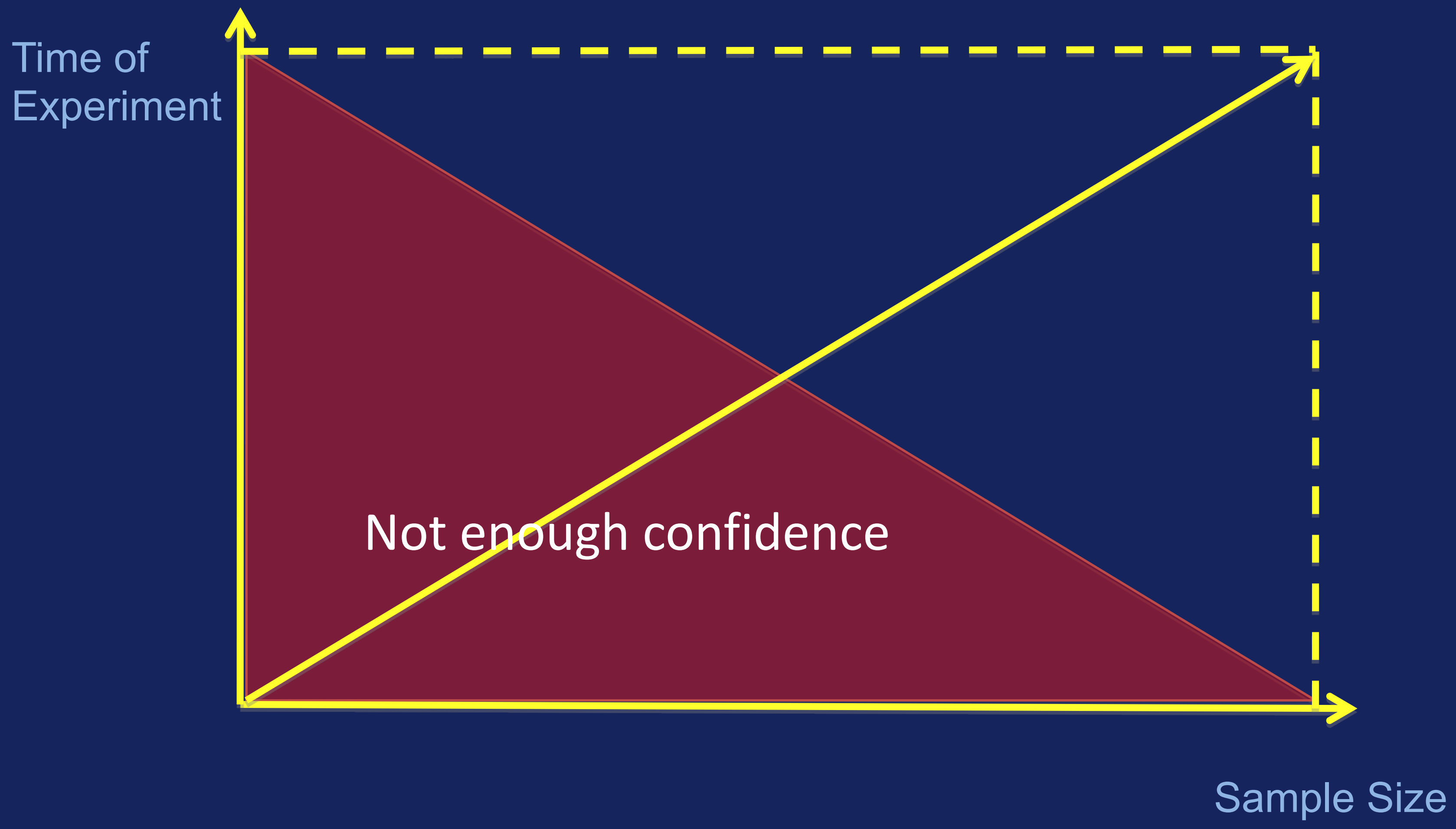


Time of Experiment
Increase Power to Detect Difference
Increase the Costs of the Evaluation
Sample Size

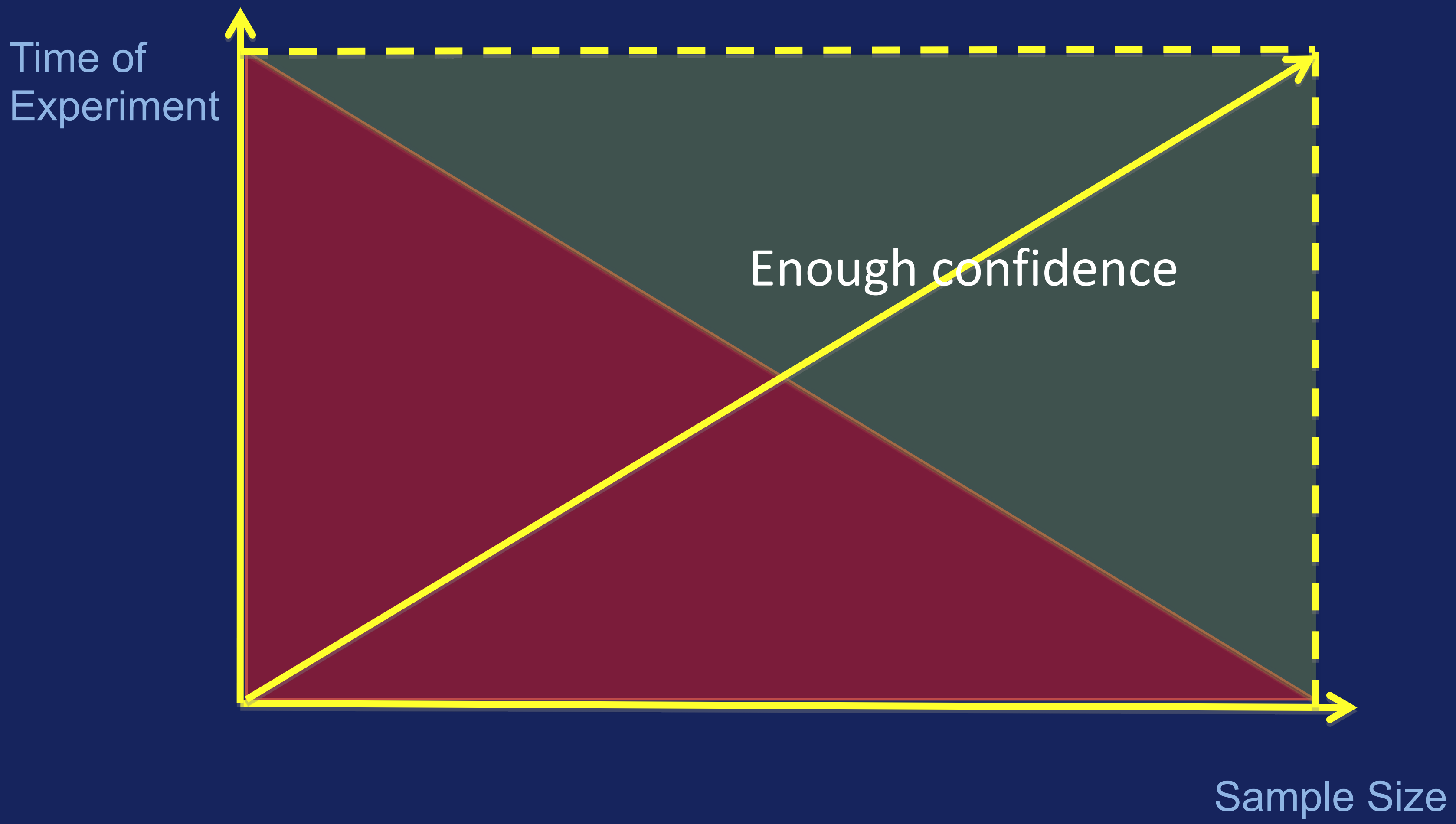
The problem

- In principle:
  - The minimum sample size
  - The minimum observational time
  - The maximum power
- So we are confident enough about the difference we find, at minimum cost

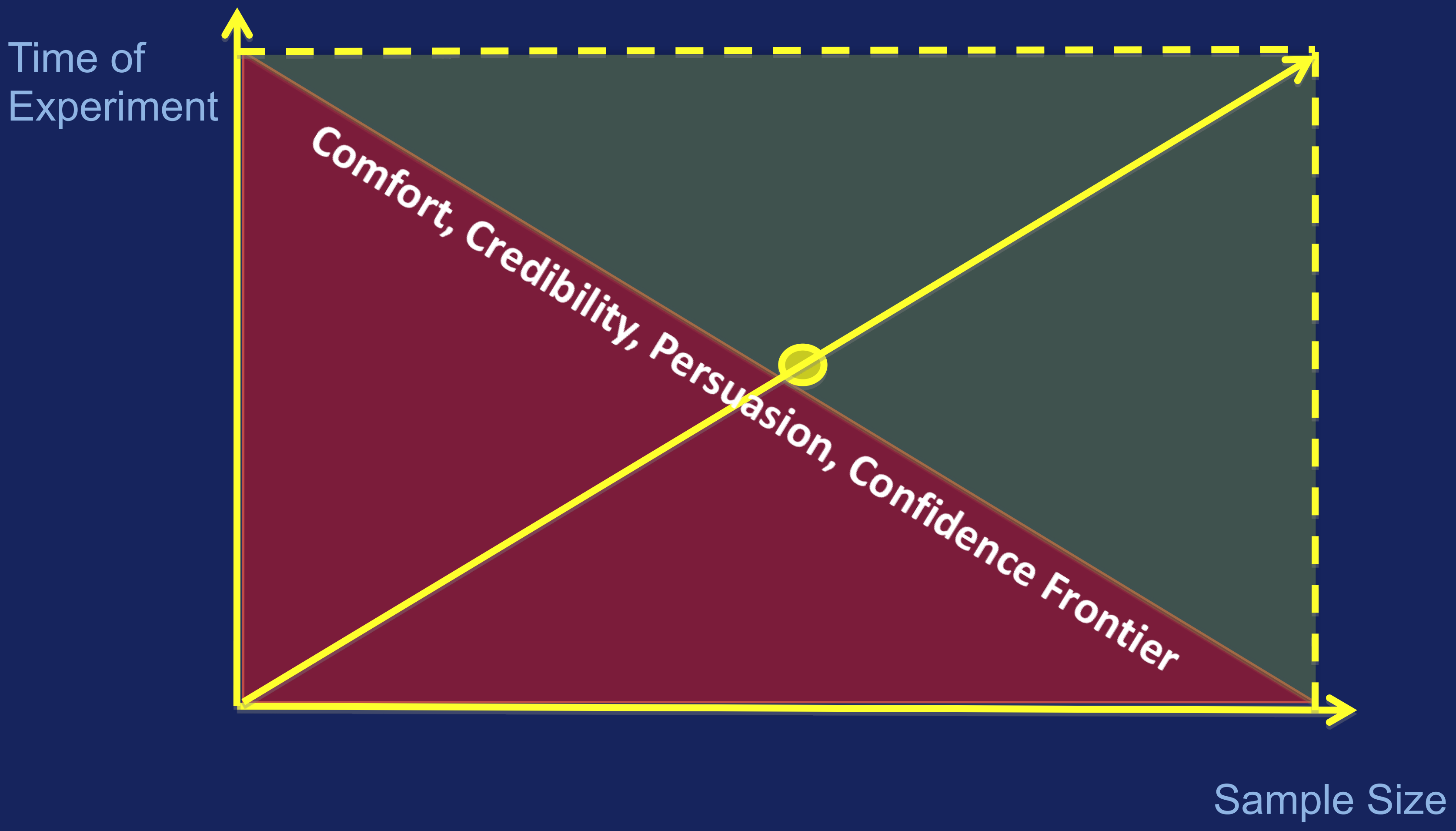
The problem



The problem

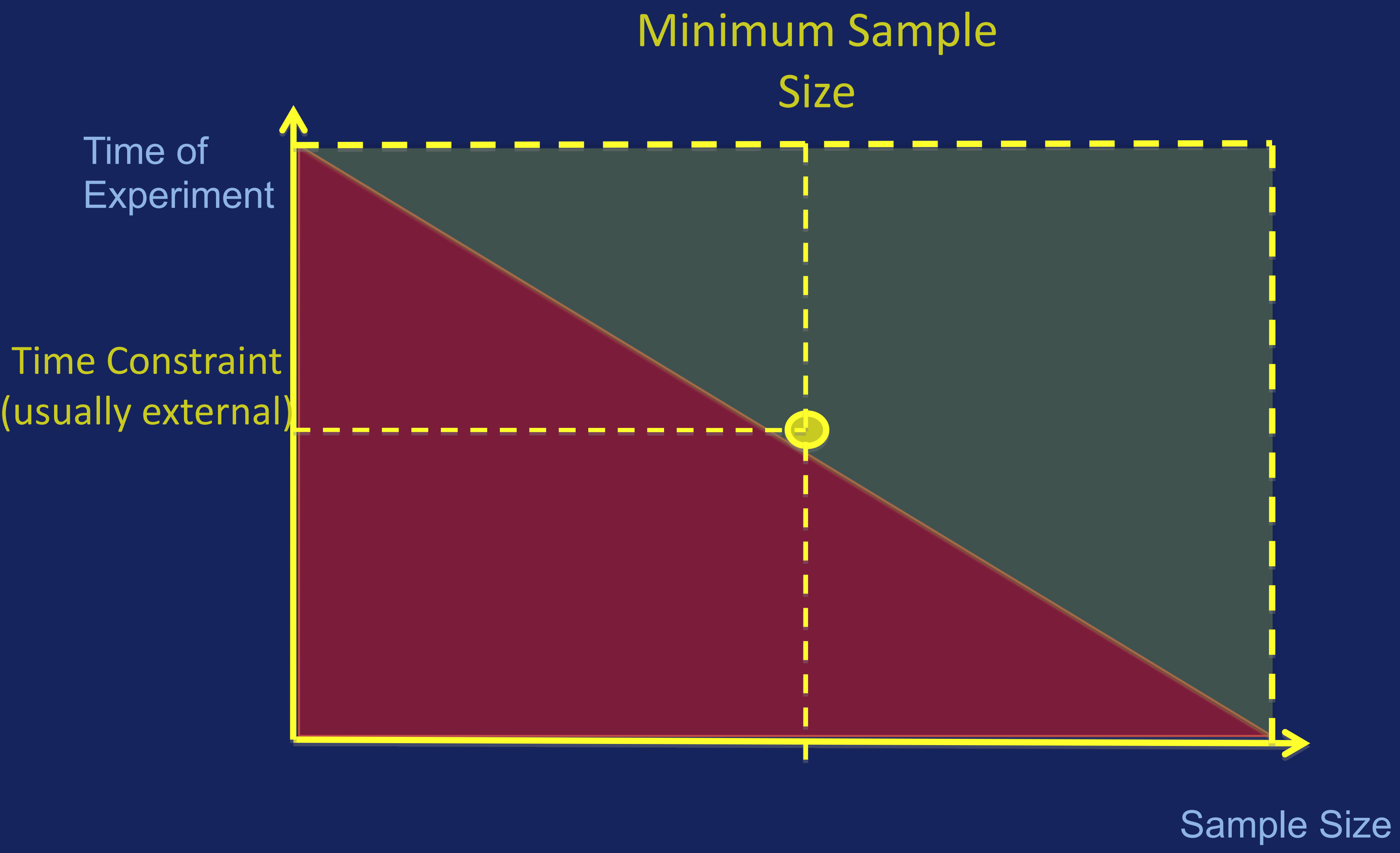


The problem



Time of Experiment
Comfort, Credibility, Persuasion, Confidence Frontier
Sample Size

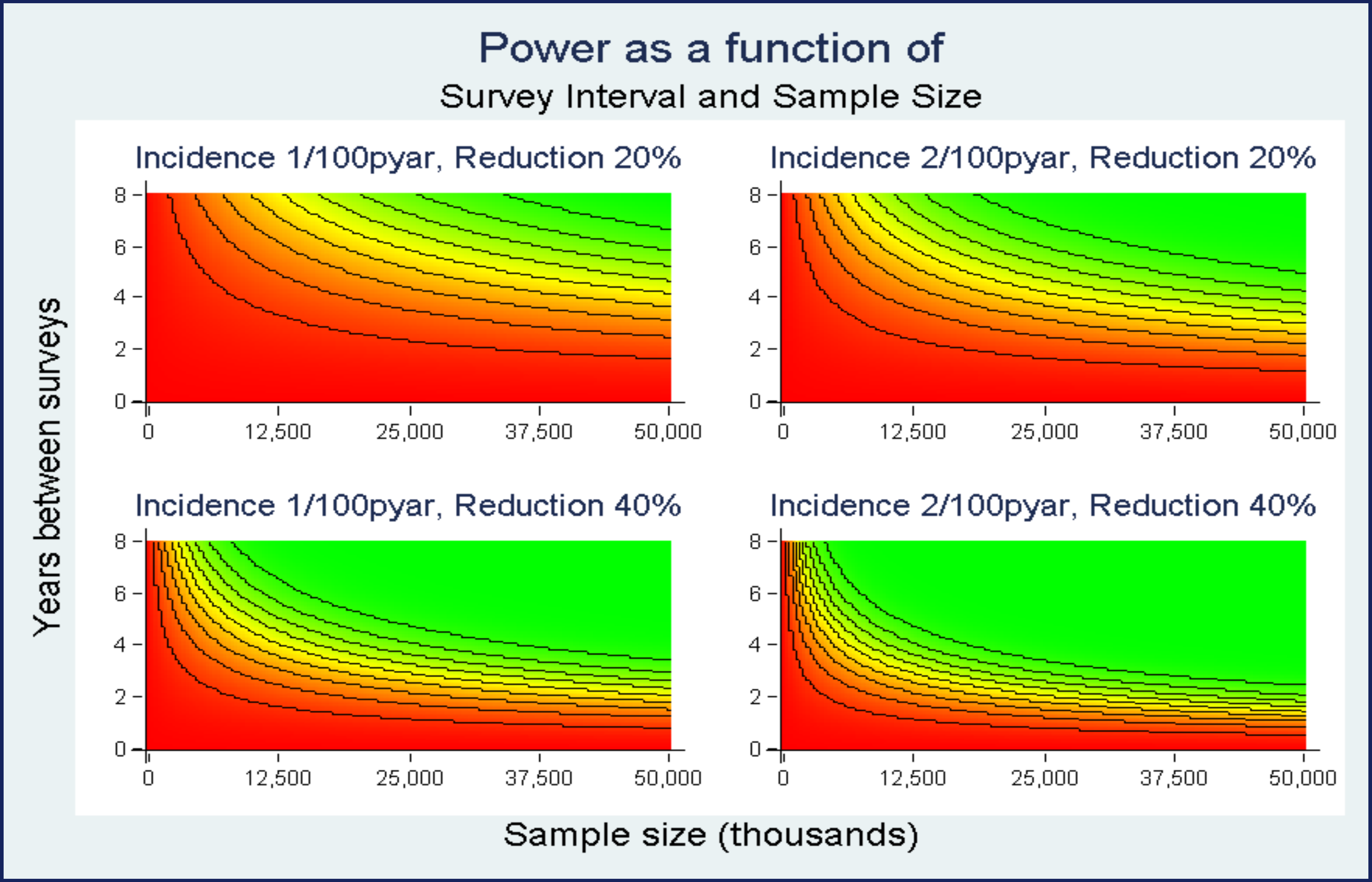
The problem


Minimum Sample
Size
Time of
Experiment
Time Constraint
(usually external)
Sample Size

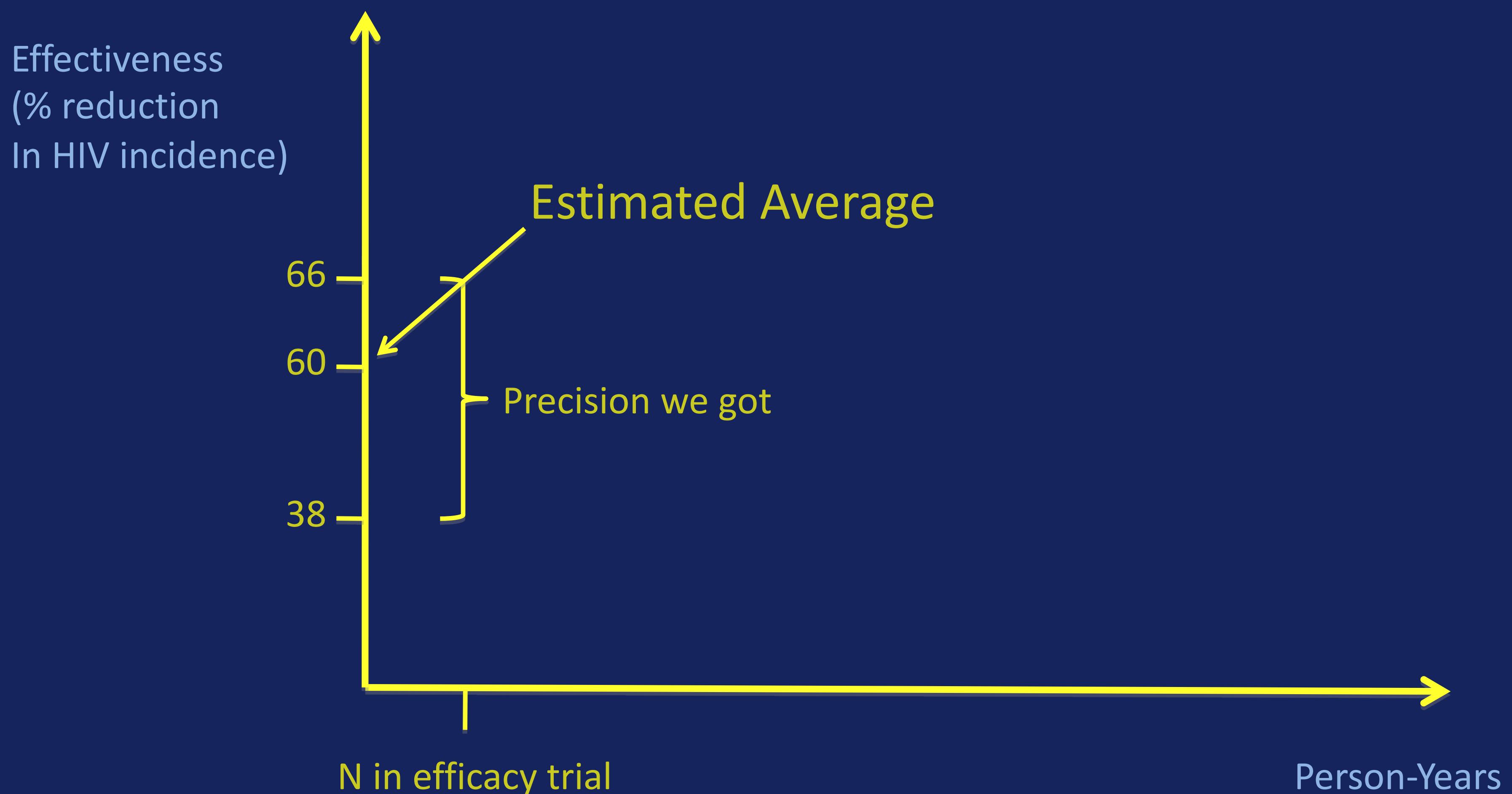
Things that increase power

- More person-years
  - More persons
  - More years
- Greater difference between control and treatment
  - More effective intervention
  - Control has no greater incidence than general population
- Clustering: Maximize independence of individuals
  - Increase number of clusters
  - Minimize intra-cluster correlation

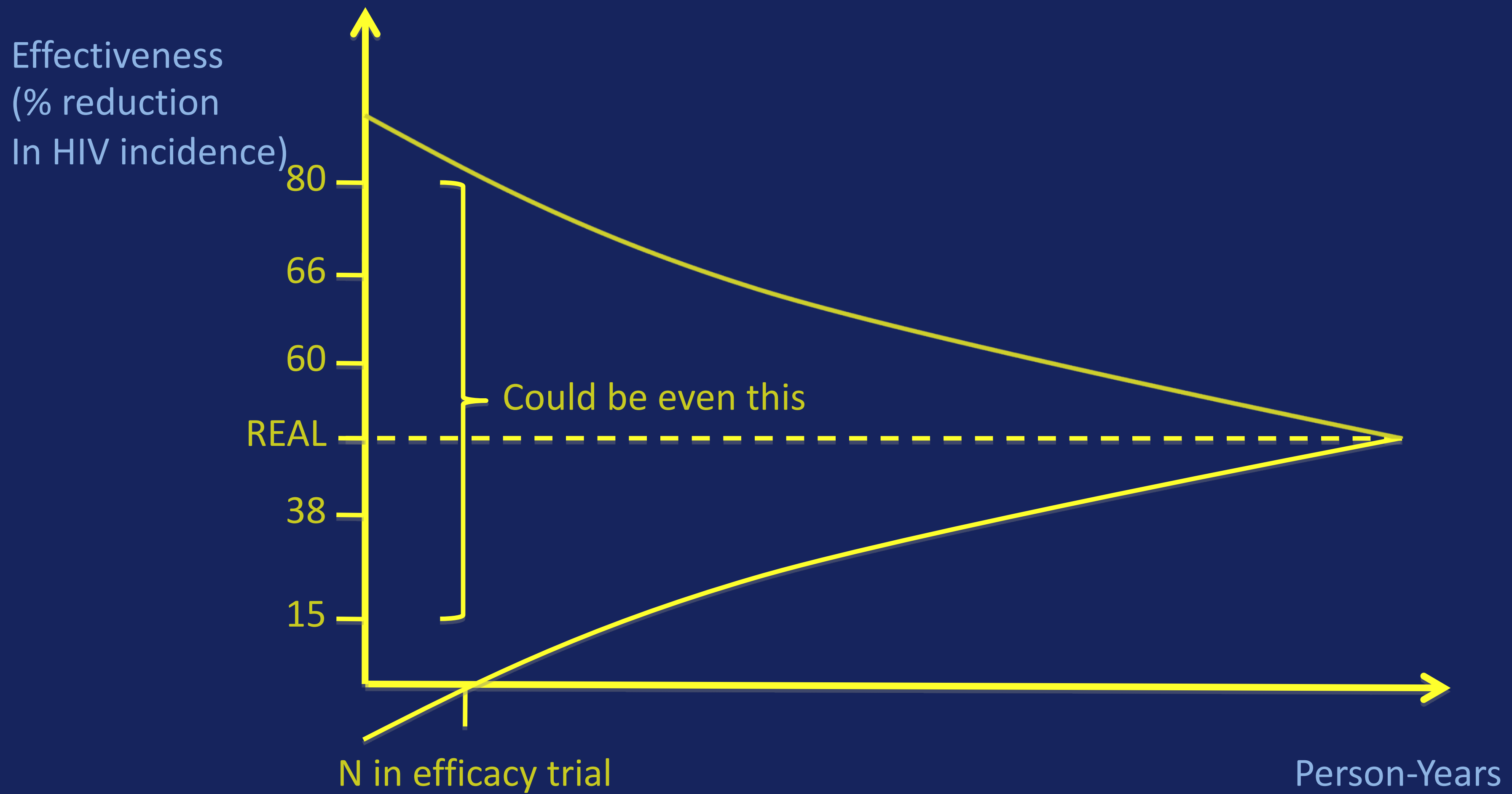
Power is higher with larger incidence in the control group or greater effect



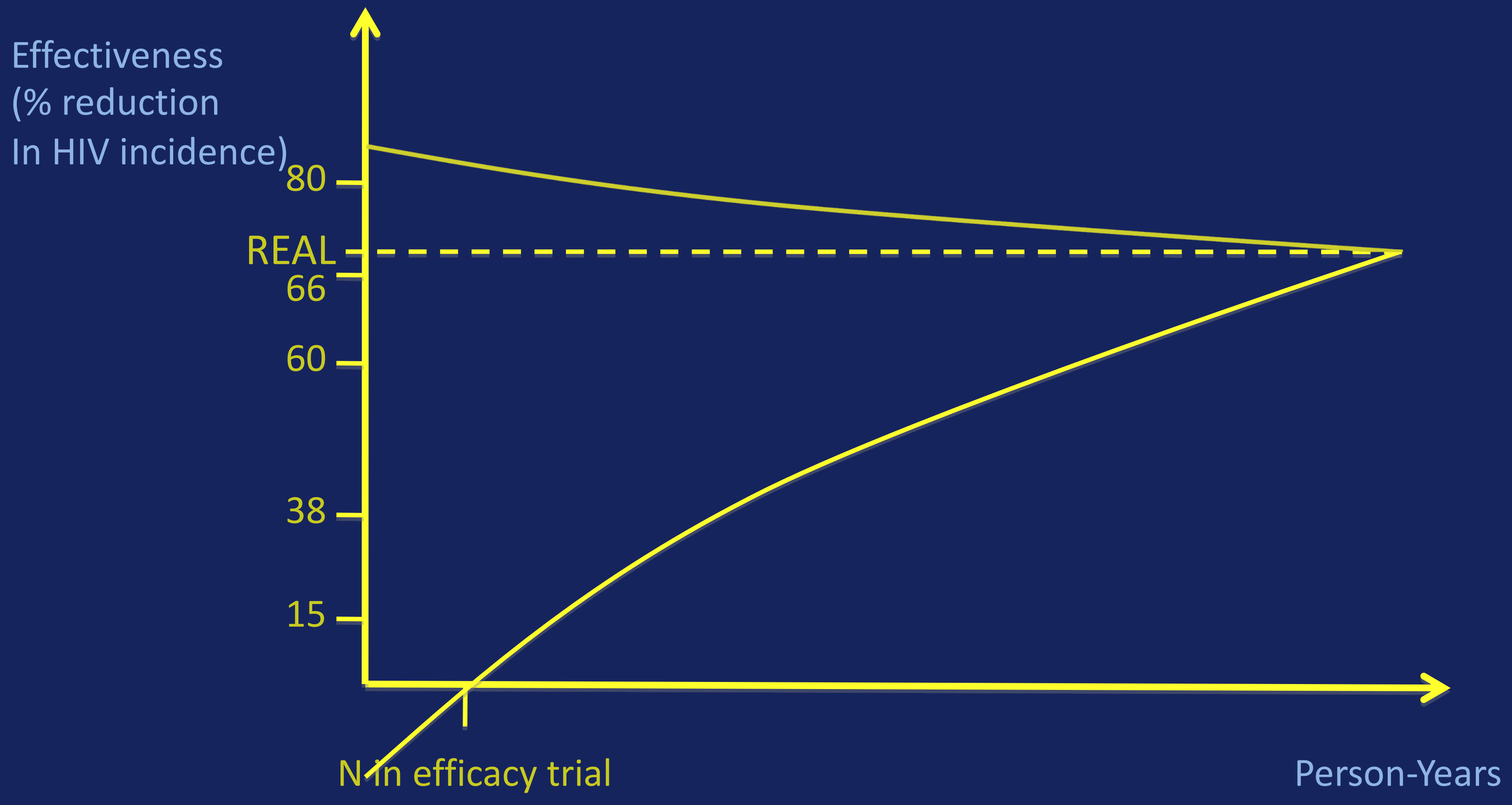
Gaining Precision



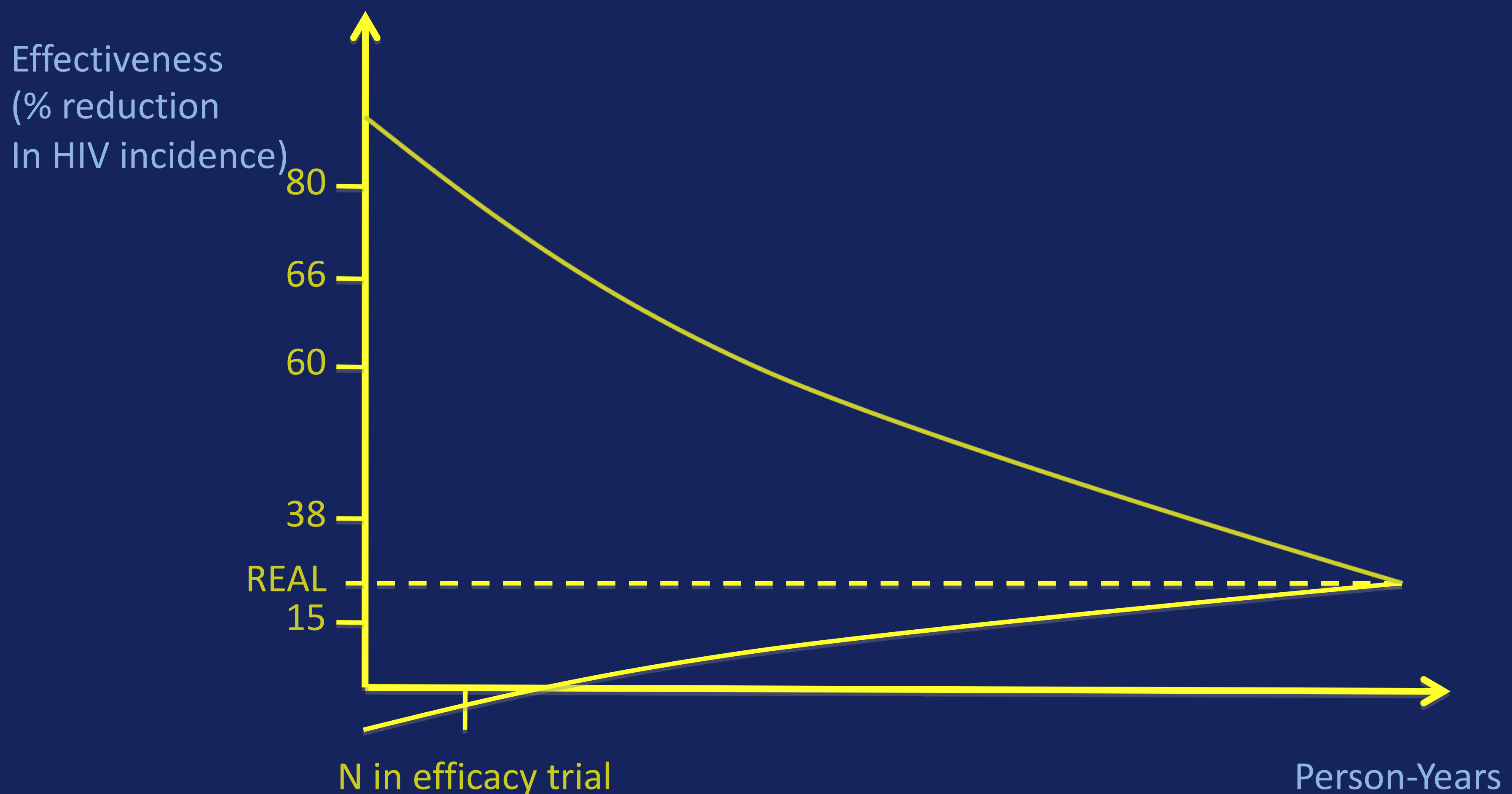
Gaining Precision



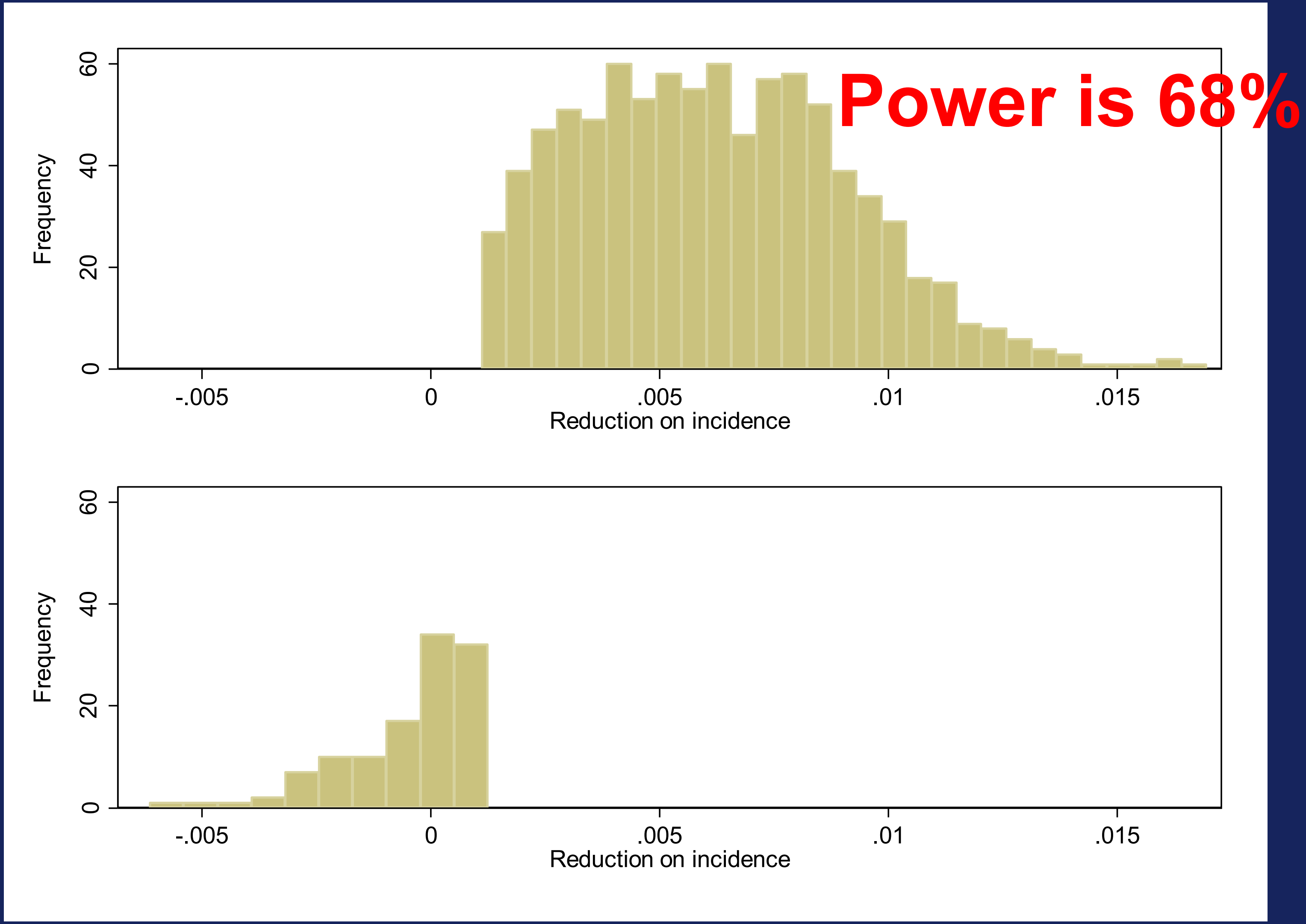
Gaining Precision



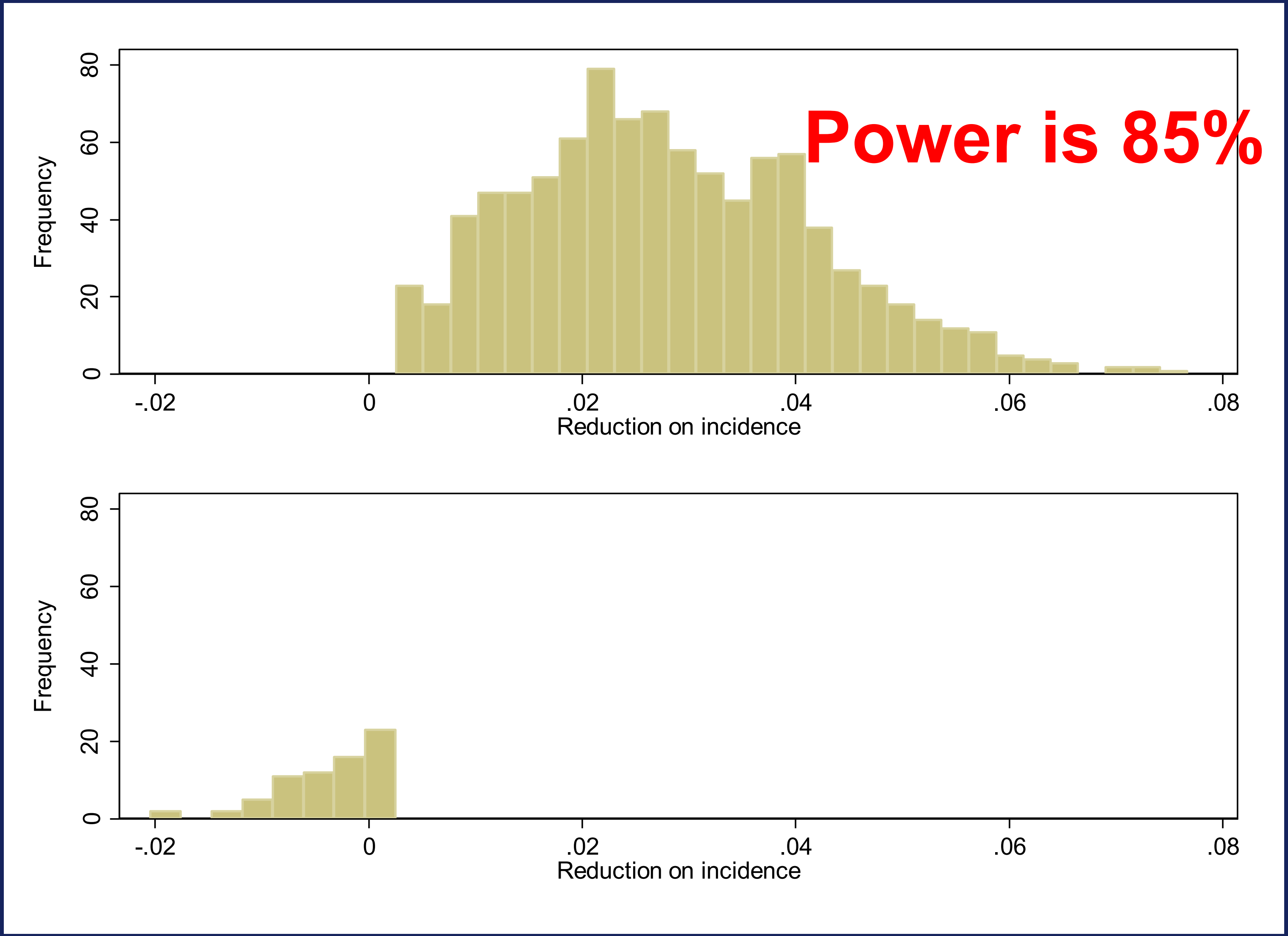
Gaining Precision



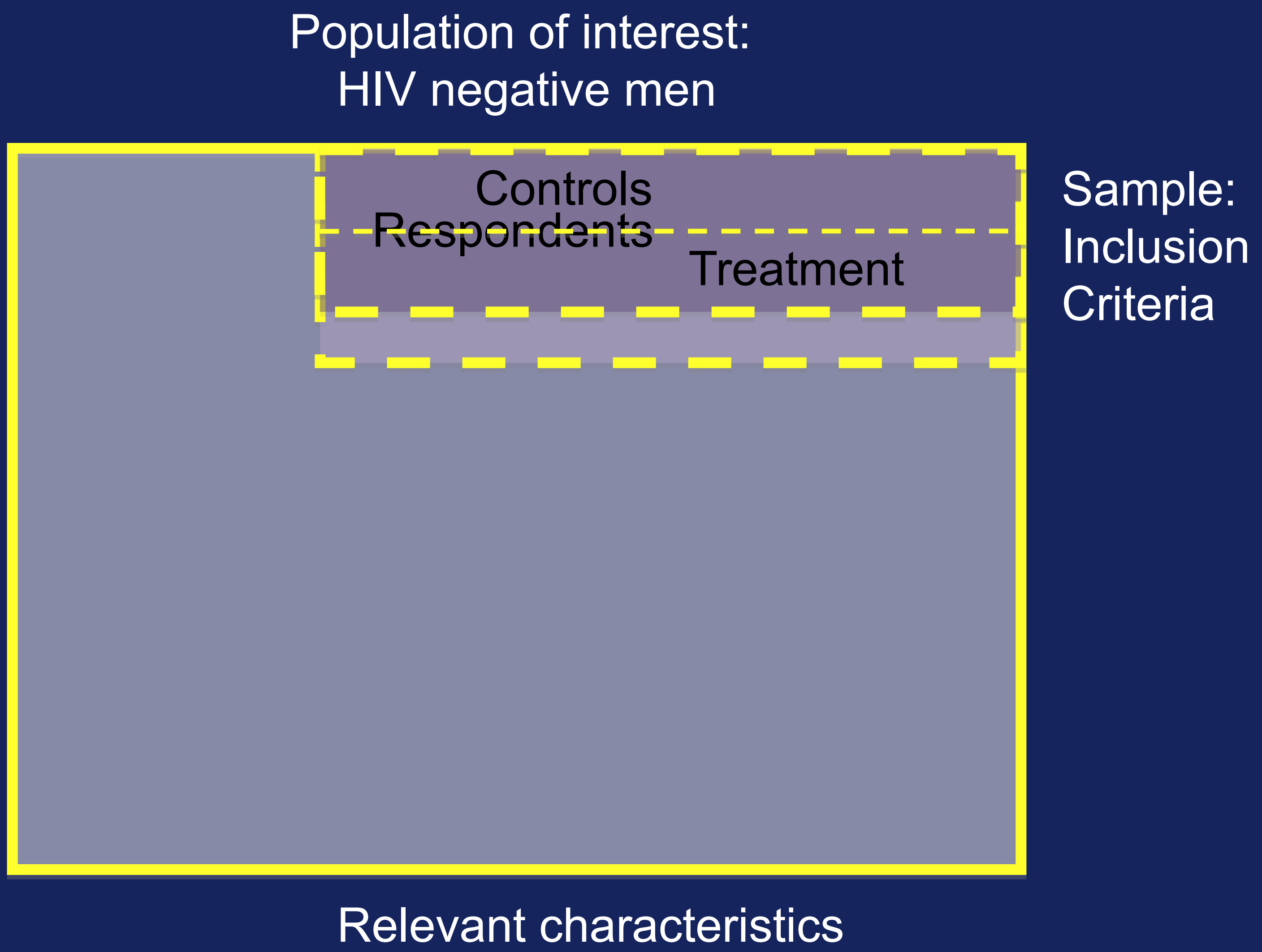
1,000 efficacy studies when the control group incidence is 1%



1,000 efficacy studies when the control group incidence is 5%



Sampling for efficacy



Effectiveness

Cost and Impact of Male Circumcision

Enter Country-specific Data

Demography
Sexual Behavior
HIV Prevalence Trends

Review or Revise Epidemiological and Economic Assumptions

Effectiveness of Male Circumcision
Epidemiological Assumptions
Economic Assumptions

Fit the Model to the Prevalence Trends

Fit the model

Set Policy Options

Specify Priority Population Groups and Target Coverage Levels
Specify Service Delivery Options

View Results

New HIV Infections
New HIV Infections by Age and Sex
HIV Incidence
Adult HIV Prevalence
Percent of Males Circumcised
Number of Circumcisions Performed
Number of Male Circumcisions per Infection Averted
Net Cost of Male Circumcisions
Net Cost and Savings per Infection Averted
AIDS Deaths
Sensitivity Analysis

Review Methods and Model Equations

Methods

About the Male Circumcision Model

This model is intended to support policy development and planning for scaling up the provision of male circumcision services. It enables analysts and decisionmakers to understand the costs and impacts of policy options. It is a part of a larger tool kit developed by the United Nations Program on HIV/AIDS (UNAIDS) and World Health Organization (WHO) that provides guidelines on comprehensive approaches to male circumcision, including types of surgical

The key policy areas addressed by the model are the following:

1. Target populations
   - Adults, young adults, adolescents, newborns, most-at-risk groups (all male)
2. Target coverage levels and rates of scale-up
3. Service delivery modes
   - Service delivery modes (hospital, clinic, mobile van; public, private, nongovernmental organization, and "other")
4. Task shifting

The costing data should already be prepared in the companion costing spreadsheets. When opening this model, be sure to choose "Update" to update the costing data.

For more information, please contact:
Health Policy Initiative, Task Order 1
Futures Group
One Thomas Circle, NW, Suite 200
Washington, DC 20005 USA
Tel: (202) 775-9680
Fax: (202) 775-9694
http://www.healthpolicyinitiative.com
policyinfo@futuresgroup.com

USAID Office of HIV/AIDS
Technical Leadership and Research Division
Emmanuel F. Njeuhmeli, MD, MPH, MBA

Effectiveness of male circumcision

Return to Menu

Reduction in annual probability of infection when circumcised		
Female to male		
- General population	60%	Sources: 1, 2, 3
- Most-at-risk population	69%	Source: 4
- Other population 1	69%	
- Other population 2	69%	
- Other population 3	69%	
Male to female	0%	

Sources:

1. Auvert B, Taljaard D, Lagarde E, Sobngwi-Tambekou J, Sitta R, et al. (2005) Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: The ANRS 1265 tr PLoS Med 2(11): e298.
Protective effect = 60% (32%-76%)
2. Gray RH, Kigozi G, Serwadda D, Makumbi F, Watya S, Nalugoda N et al. Male circumcision for HIV prevention n men in Rakai, Uganda: a randomized trial. Lancet 2007; 369: 767-66.
Protective effect = 55% (22%-75%)
3. Bailey RC, Moses S, Parker CB, Agot K, Maclean I, Krieger JN, et al. Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomized controlled trial. Lancet 2007; 369: 643-56.
Protective effect = 60% (32%-77%)
4. Weiss HA, Quigley MA, Hayes R. Male circumcision and risk of HIV infection in sub-Saharan Afri a systematic review and meta-analysis. AIDS 2000, 14:2361-2370.
Protective effect among most-at-risk men = 71% (59%-80%)

Calculated force of infection for those not circumcised (For information only, no inputs required)	
Male 15-29	0.13
Female 15-29	2.05
Male 30-49	0.92
Female 30-49	2.27

Effectiveness of male circumcision

Return to Menu

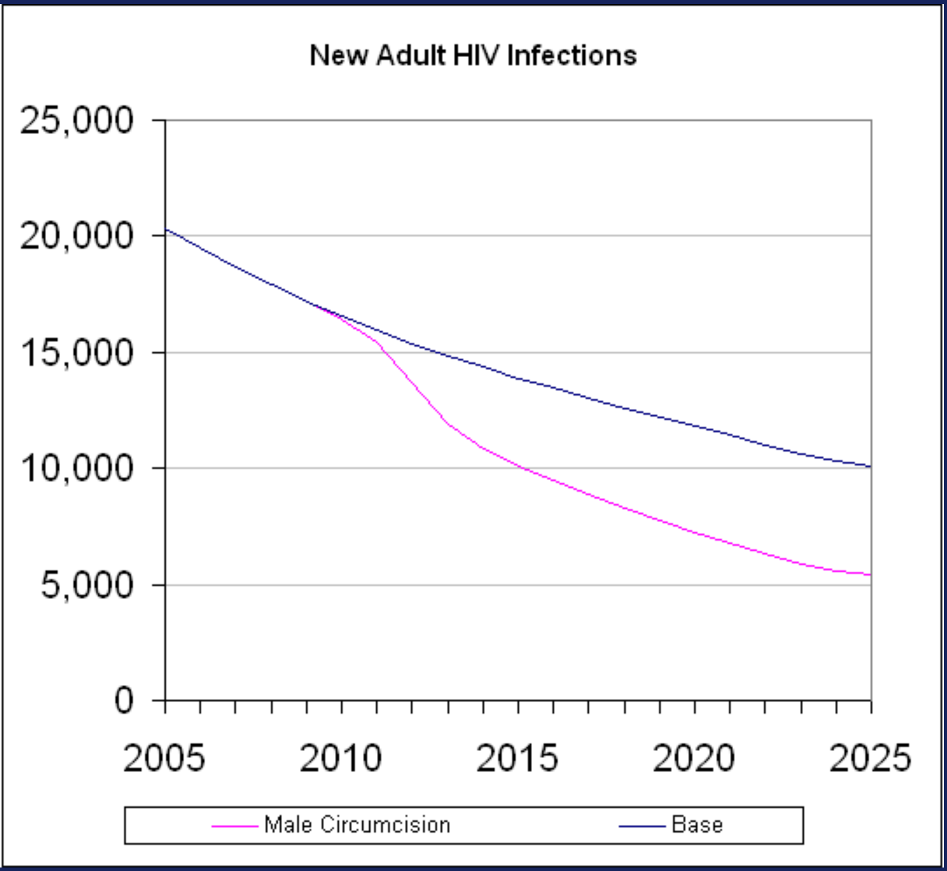
Reduction in annual probability of infection when circumcised		
Female to male		
- General population	20%	Sources: 1, 2, 3
- Most-at-risk population	20%	Source: 4
- Other population 1	20%	
- Other population 2	20%	
- Other population 3	20%	
Male to female	0%	

Sources:

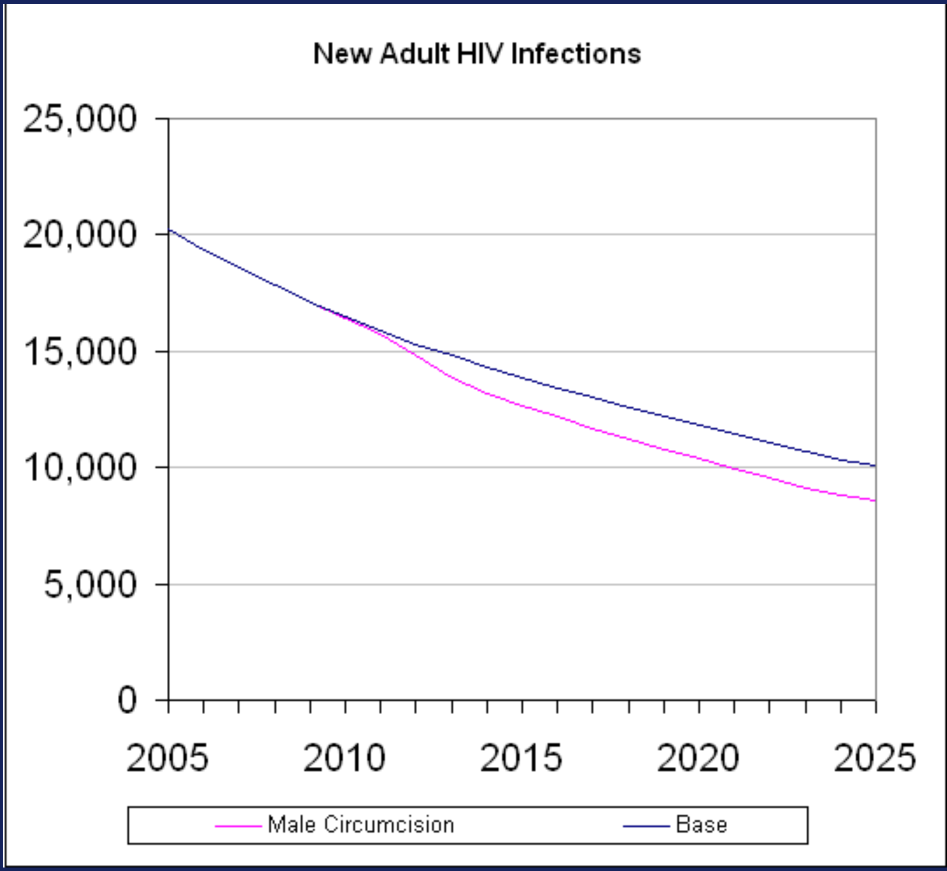
1. Auvert B, Taljaard D, Lagarde E, Sobngwi-Tambekou J, Sitta R, et al. (2005) Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: The ANRS 1265 tri PLoS Med 2(11): e298.
Protective effect = 60% (32%-76%)
2. Gray RH, Kigozi G, Serwadda D, Makumbi F, Watya S, Nalugoda N et al. Male circumcision for HIV prevention n men in Rakai, Uganda: a randomized trial. Lancet 2007; 369: 767-66.
Protective effect = 55% (22%-75%)
3. Bailey RC, Moses S, Parker CB, Agot K, Maclean I, Krieger JN, et al. Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomized controlled trial. Lancet 2007; 369: 643-56.
Protective effect = 60% (32%-77%)
4. Weiss HA, Quigley MA, Hayes R. Male circumcision and risk of HIV infection in sub-Saharan Afric a systematic review and meta-analysis. AIDS 2000, 14:2361-2370.
Protective effect among most-at-risk men = 71% (59%-80%)

Calculated force of infection for those not circumcised (For information only, no inputs required)	
Male 15-29	0.13
Female 15-29	2.05
Male 30-49	0.92
Female 30-49	2.27

Less effective intervention - less impact

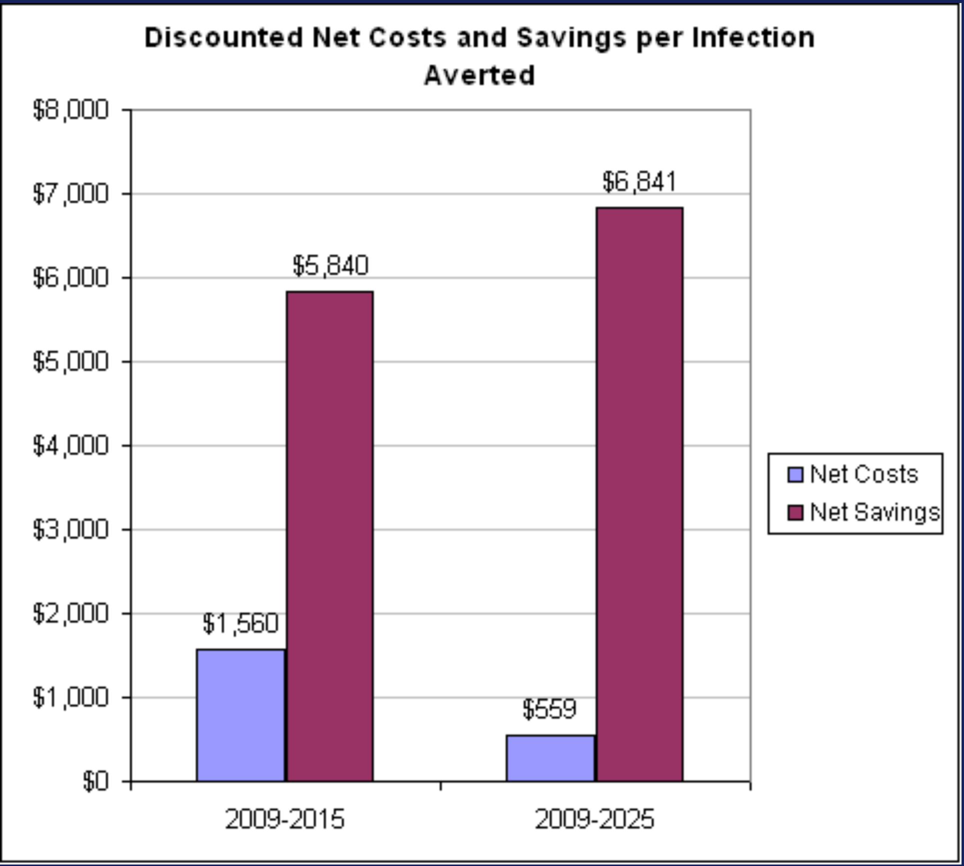


60% Effectiveness

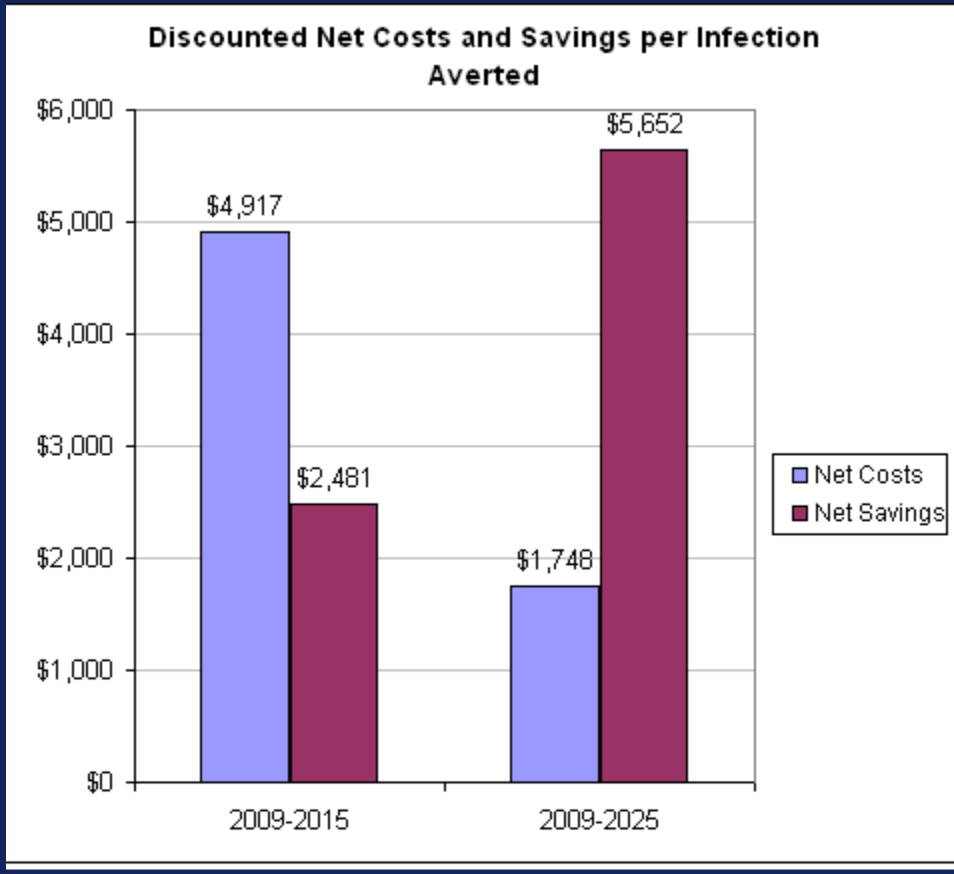


20% Effectiveness

Less effective intervention - less cost-effective



60% Effectiveness



20% Effectiveness

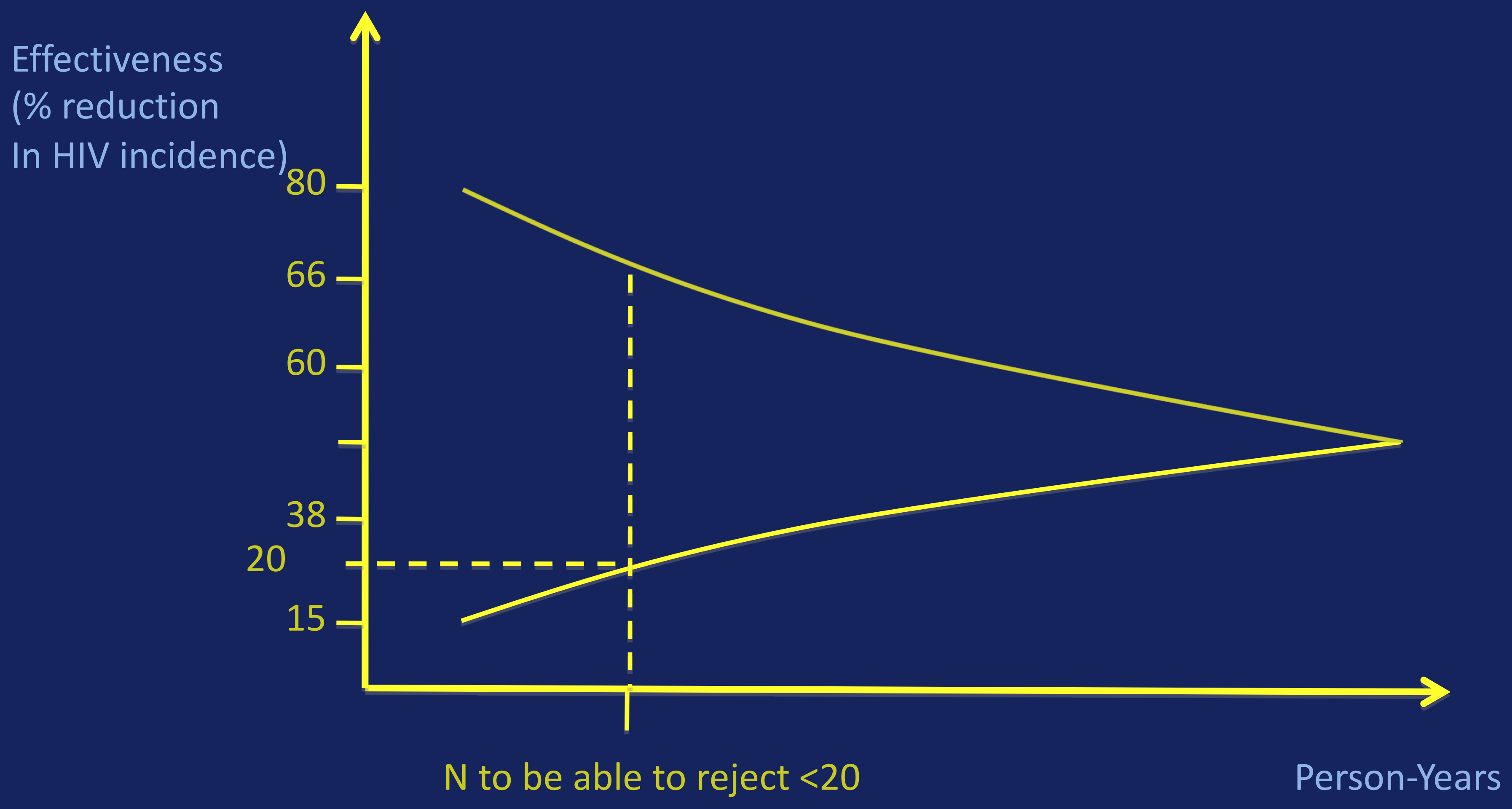
Hypothesis to reject

- Circumcision has no impact
- **Circumcision has too little impact**
- **Intensive Circumcision Program has no more impact than Routine Circumcision Program**
- **Circumcision has no benefit for women**

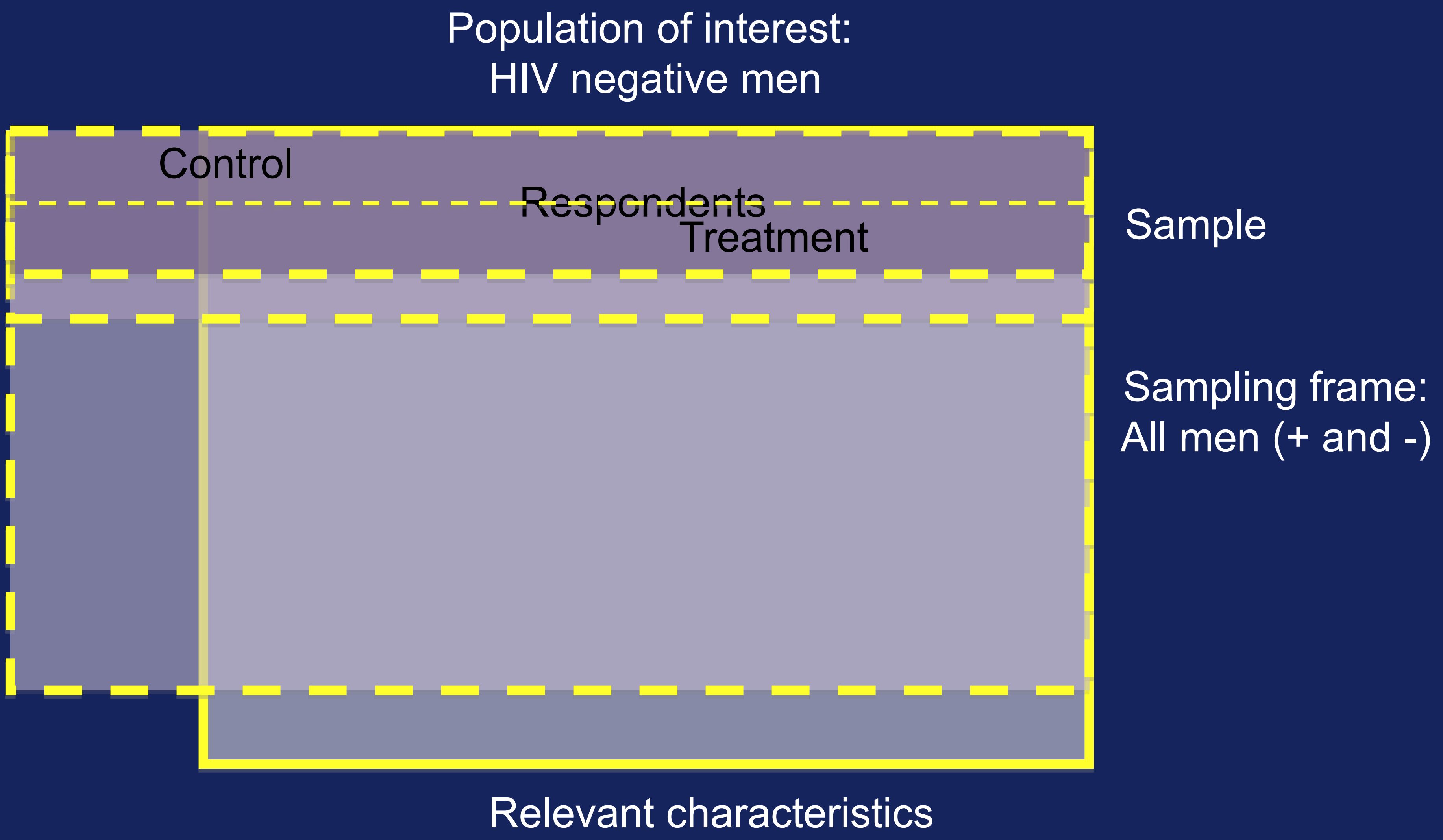
Differences between effectiveness and efficacy that affect sampling

- Main effect on HIV incidence in HIV- men
  - Null hypothesis: impact > 0 (+)
  - Effect size because of standard of care (+)
- Investigate determinants of effectiveness
  - Supply side (+ / -)
  - Demand side (+ / -)
- Investigate impact on secondary outcomes and their determinants (+ / -)
- Seek “external validity” on effectiveness issues

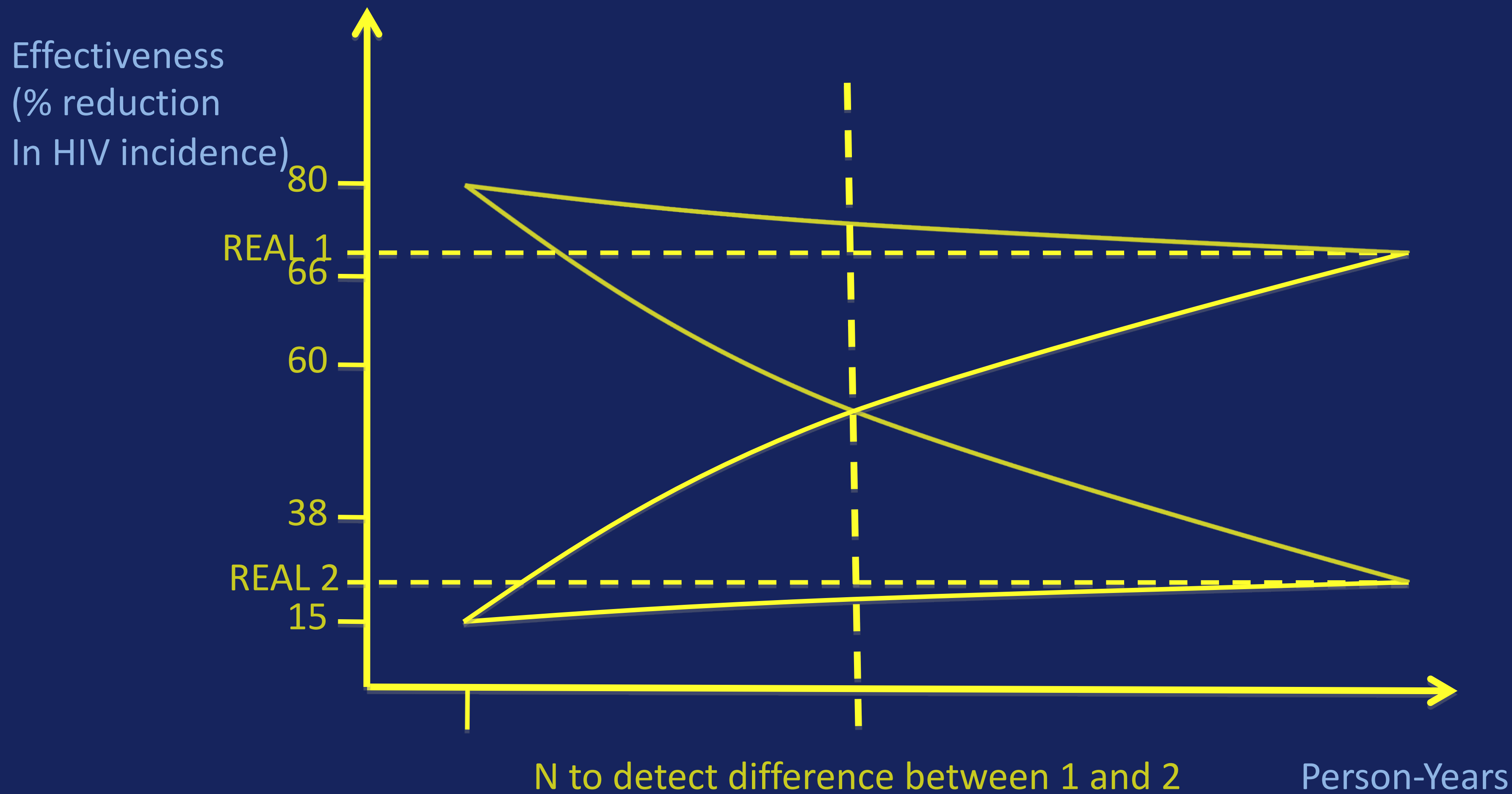
Power increases with sample size



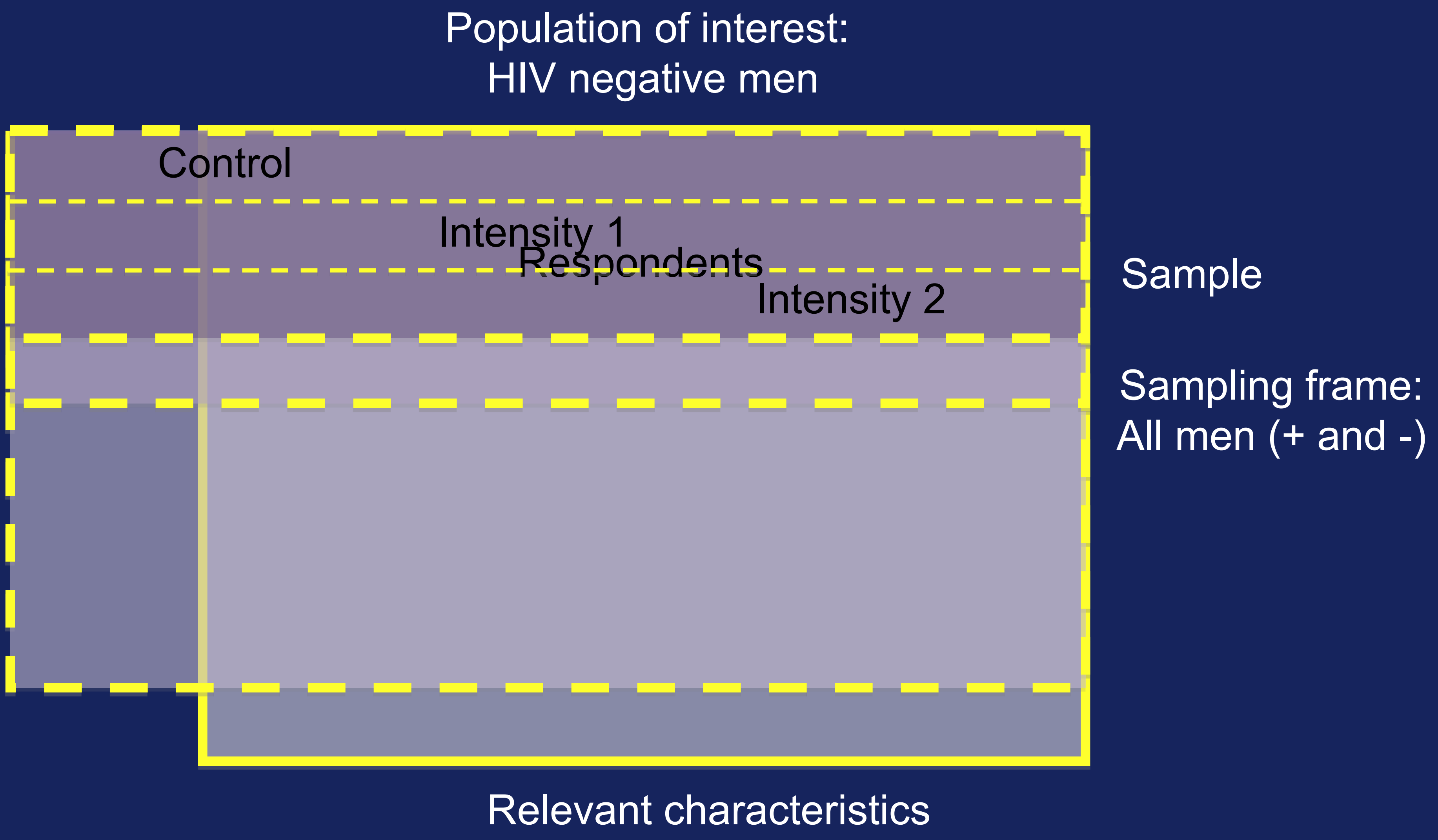
Sampling for effectiveness



Two levels of effectiveness



Sampling for effectiveness



Sampling methods for effectiveness

- Probability sampling
  - Simple random: each unit in the sampling frame has the same probability of being selected into the sample
  - Stratified: first divide the sampling frame into strata (groups, blocks), then do a simple random sample within each strata
  - Clustered: sample clusters of units. Eg. villages with all the persons that live there
    - One stage: Random sample of villages, then survey all men in selected villages
    - Two stage: Random sample of villages, then random sample of men in selected villages

Cluster Sampling Design

Cluster Sampling

- In some situations, individual random samples are not feasible
  - When interventions are delivered at the facility/community level
  - When constructing a frame of the observation units may be difficult, expensive, or even impossible
    - Customers of a store
    - Birds in a region
  - When is of interest to identify community level impact
  - When budget constraints don’t allow it

M.K. Campbell et al. Computers in Biology and Medicine 34 (2004) 113 – 125

Clustering and sample size

- Clustering reduces efficiency of the design
  - Standard sample size calculation for individual-based studies only accommodate for variation between individuals
  - In cluster studies, there are two components of variation
    - Variation among individuals within clusters
    - Variation in outcome between clusters

Clustering and sample size

- Individual-based studies assume independence of outcomes among individuals
- In cluster randomization:
  - Individuals within a cluster are more likely to be similar
- Measure of this intracluster dependence among individuals is ICC
  - Based in within-cluster variance
    - High when individuals in cluster are more “similar”
- Not taking ICC into account may lead to under-powered study (too small sample)

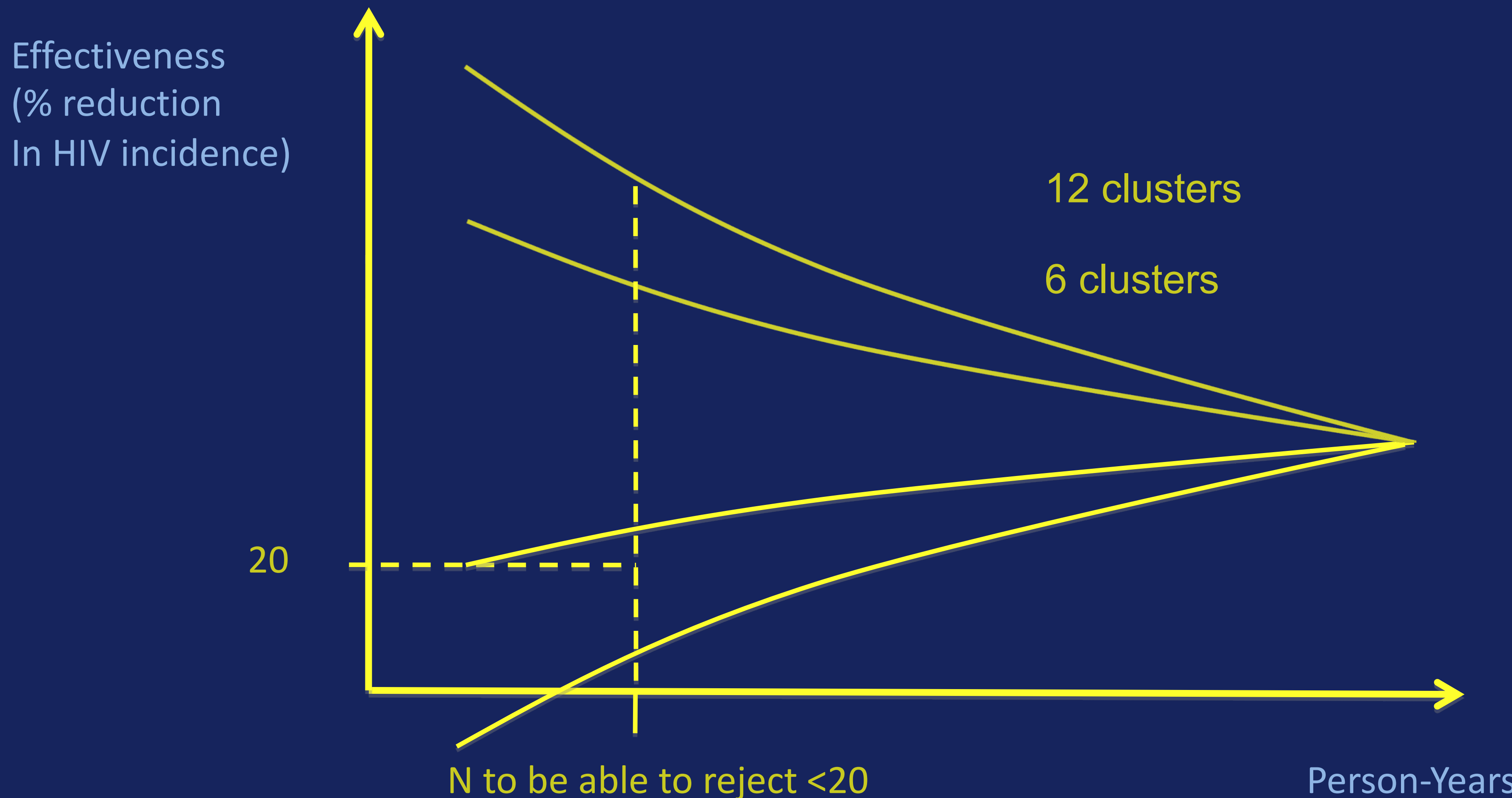
Taking ICC into account

- In a cluster randomized design, in order to achieve the equivalent power of a individual random study, sample size require to be inflated by a factor:

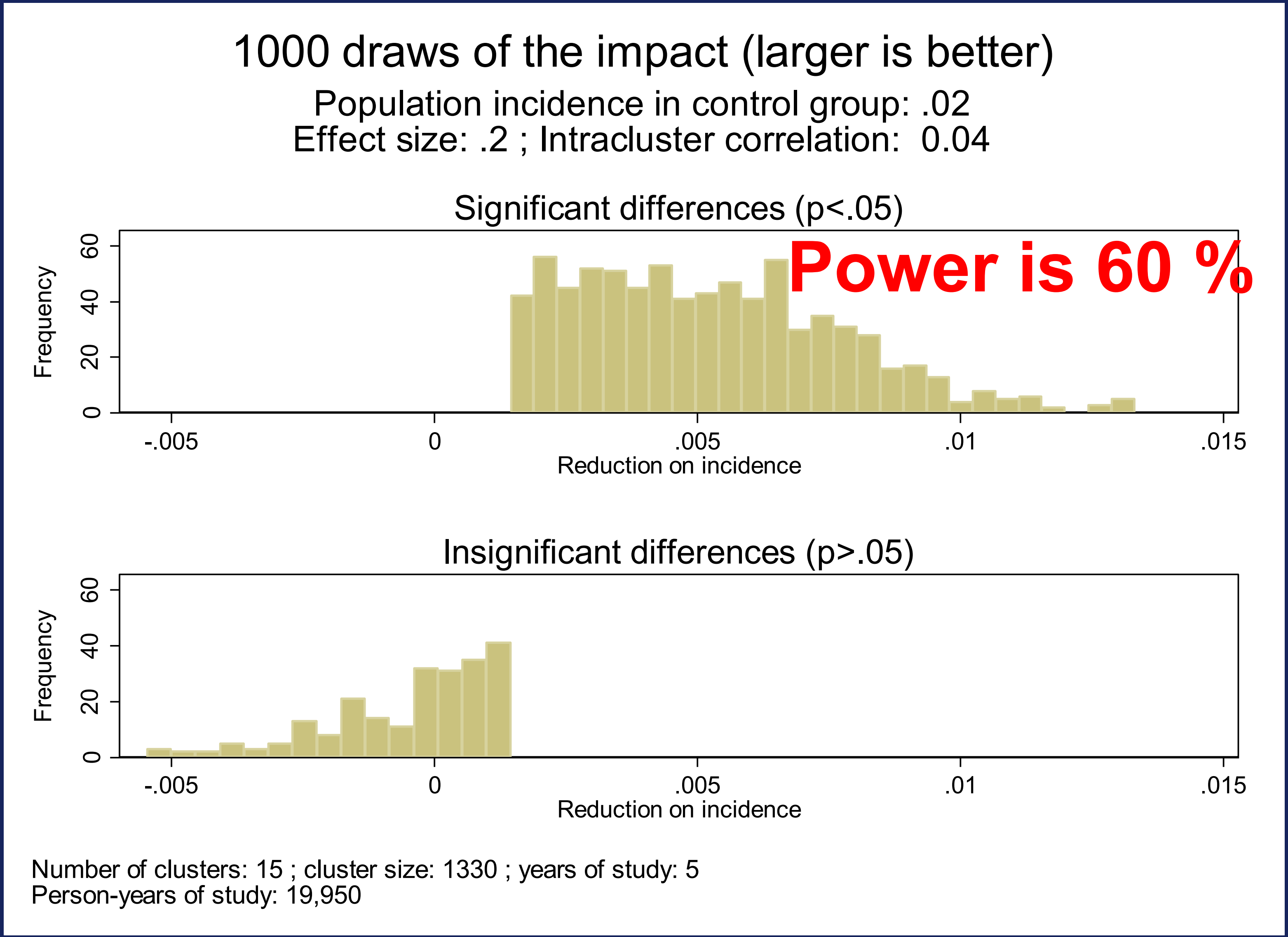
  $$Deff = 1 + (\tilde{n} - 1)\,\rho$$

  to consider cluster effect
- $\tilde{n}$ = average cluster size
- ρ = ICC
- Assuming clusters of similar size

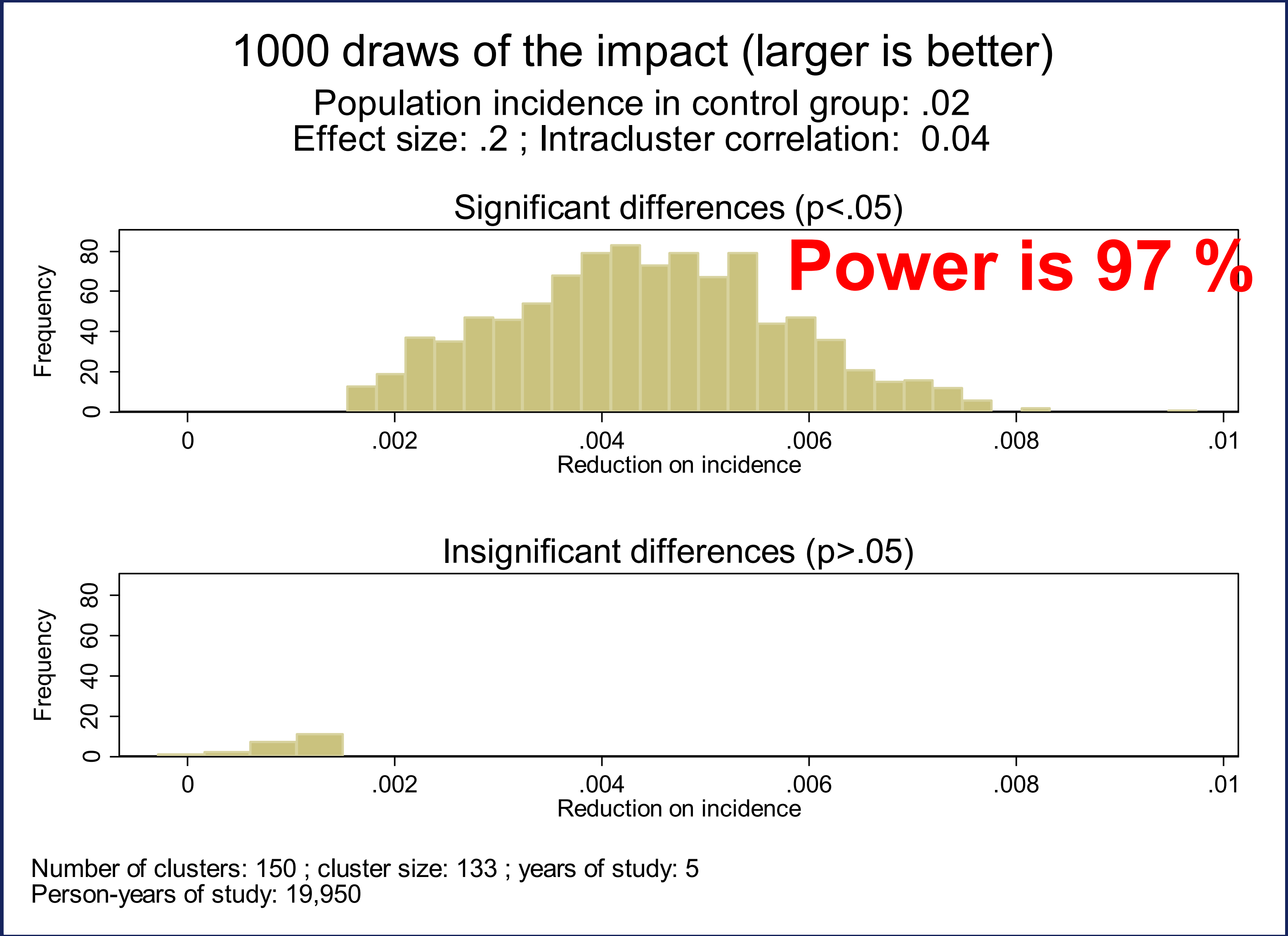
How big is the impact of cluster design on sample size



When 19,950 individuals are in 15 clusters



When 19,950 individuals are in 150 clusters



Increasing sample size

- Increasing the number of clusters vs increasing the number the individuals per cluster
- Increasing the number of clusters has a much stronger effect on power and confidence
  – Intuitively, the sample is the number of units (clusters) at the level where the random assignment takes place. It is not the same as the number of people surveyed
- Challenge is to engineer the logistics to to maximize the number of clusters, given

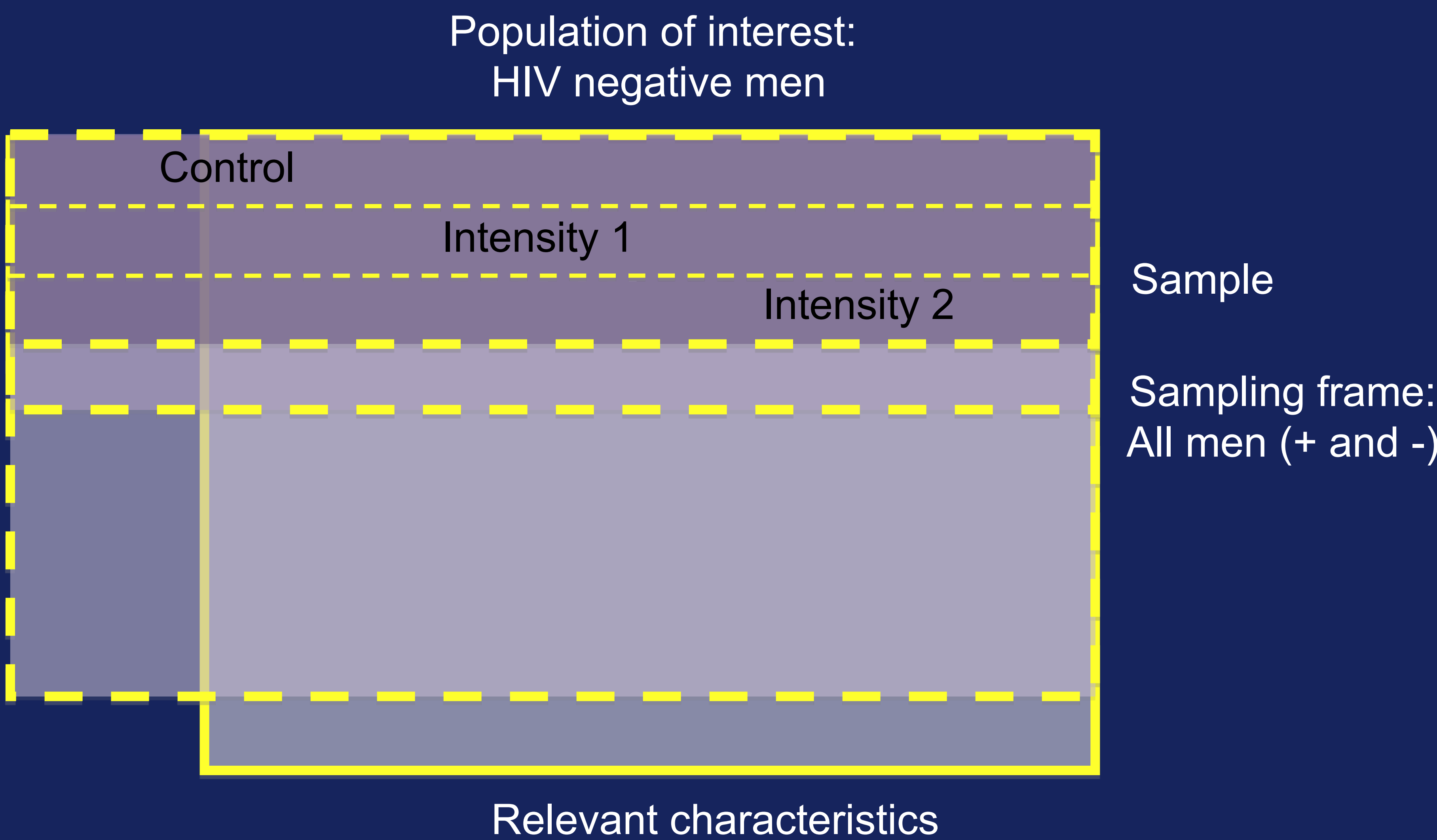
How big is the impact of cluster design on sample size

Impact on costs?

Things that affect costs

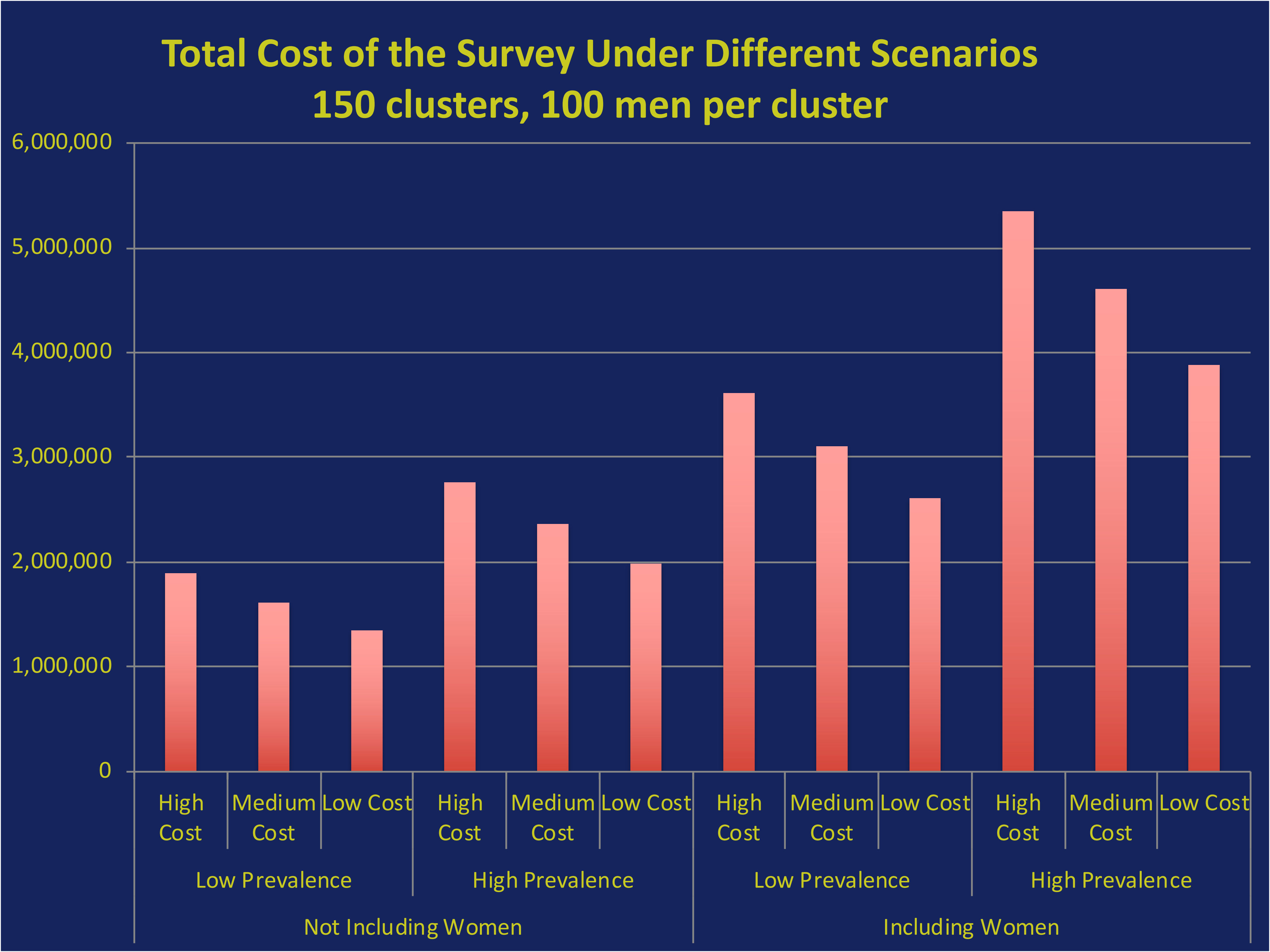
- Including HIV positive men
- Including women
- Prevalence of HIV
- Length of questionnaire
  - To measure more outcomes
  - To measure implementation of intervention and costs
    - For cost-effectiveness
    - For control quality and other characteristics of the intervention

Sampling for effectiveness



Some Scenarios

- 150 clusters, 100 men per cluster
- Including women → double number of HIV tests
- Low and High prevalence → additional men to be surveyed
- High, medium, low cost
  - Dispersion of clusters → distance among them
  - Length of questionnaire → time in fieldwork, data collection staff


Total Cost of the Survey Under Different Scenarios
150 clusters, 100 men per cluster
6,000,000
5,000,000
4,000,000
3,000,000
2,000,000
1,000,000
0
High Cost
Medium Cost
Low Cost
High Cost
Medium Cost
Low Cost
High Cost
Medium Cost
Low Cost
High Cost
Medium Cost
Low Cost
Low Prevalence
High Prevalence
Low Prevalence
High Prevalence
Not Including Women
Including Women

Conclusions

- Philosophy of sample design is different for efficacy and effectiveness studies
  - Efficacy: narrow & deep
  - Effectiveness: broad & shallow
- Many of the special requirements of effectiveness sampling will increase sample size
- Clustering reduces data collection costs but at a sacrifice of power
- Survey costs also affected
  - Number of indicators collected
  - Number of non-index cases interviewed
- Most cost-effective way to reject your “hated hypothesis” is through randomized, efficeintly powered, sampling