The analysis of active-control trials: missing the point

David Dunn
MRC Clinical Trials Unit at UCL
d.dunn@ucl.ac.uk

Active-control trials

- Experimental treatment is compared with an established treatment
- Performed when the inclusion of a placebo control group is deemed to be unethical
- Often designed as non-inferiority trials
- For time-to-event outcomes, the standard primary metric is: rate ratio, hazard ratio, or rate difference

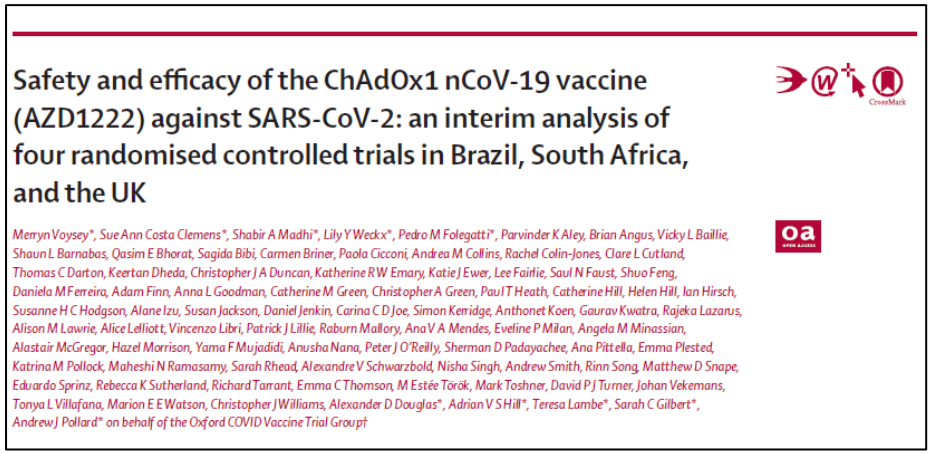
Structure

- Three examples where standard analytical approach is misleading
- Introduce new metric
- Re-analyse the three examples
- Challenges in estimating new metric
- Sample size
- Conclusions

Vaccine efficacy against COVID-19

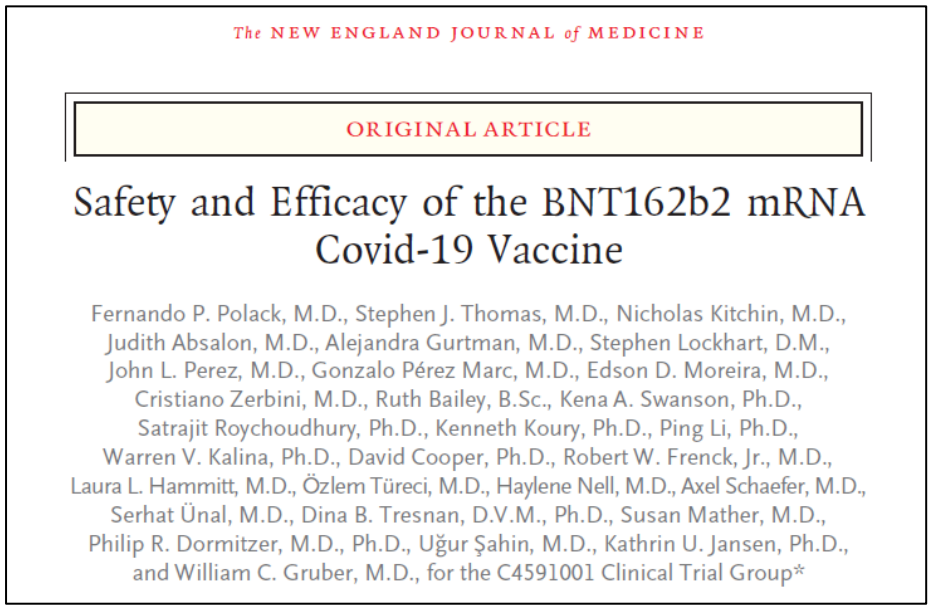
~80%
(if prime-boost interval > 12 weeks)

~95%



Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK

Merryn Voysey*, Sue Ann Costa Clemens*, Shabir A Madhi*, Lily Y Weckx*, Pedro M Folegatti*, Parvinder K Aley, Brian Angus, Vicky L Baillie, Shaun L Barnabas, Qasim E Bhorat, Sagida Bibi, Carmen Briner, Paola Cicconi, Andrea M Collins, Rachel Colin-Jones, Clare L Cutland, Thomas C Darton, Keertan Dheda, Christopher J A Duncan, Katherine RW Emary, Katie J Ewer, Lee Fairlie, Saul N Faust, Shuo Feng, Daniela M Ferreira, Adam Finn, Anna L Goodman, Catherine M Green, Christopher A Green, Paul T Heath, Catherine Hill, Helen Hill, Ian Hirsch, Susanne H C Hodgson, Alane Izu, Susan Jackson, Daniel Jenkin, Carina C D Joe, Simon Kerridge, Anthonet Koen, Gaurav Kwatra, Rajeka Lazarus, Alison M Lawrie, Alice Lelliott, Vincenzo Libri, Patrick J Lillie, Raburn Mallory, Ana V A Mendes, Eveline P Milan, Angela M Minassian, Alastair McGregor, Hazel Morrison, Yama F Mujadidi, Anusha Nana, Peter J O'Reilly, Sherman D Padayachee, Ana Pittella, Emma Plested, Katrina M Pollock, Maheshi N Ramasamy, Sarah Rhead, Alexandre V Schwarzbold, Nisha Singh, Andrew Smith, Rinn Song, Matthew D Snape, Eduardo Sprinz, Rebecca K Sutherland, Richard Tarrant, Emma C Thomson, M Estée Török, Mark Toshner, David P J Turner, Johan Vekemans, Tonya L Villafana, Marion E E Watson, Christopher J Williams, Alexander D Douglas*, Adrian V S Hill*, Teresa Lambe*, Sarah C Gilbert*, Andrew J Pollard* on behalf of the Oxford COVID Vaccine Trial Group†



The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine

Fernando P. Polack, M.D., Stephen J. Thomas, M.D., Nicholas Kitchin, M.D., Judith Absalon, M.D., Alejandra Gurtman, M.D., Stephen Lockhart, D.M., John L. Perez, M.D., Gonzalo Pérez Marc, M.D., Edson D. Moreira, M.D., Cristiano Zerbini, M.D., Ruth Bailey, B.Sc., Kena A. Swanson, Ph.D., Satrajit Roychoudhury, Ph.D., Kenneth Koury, Ph.D., Ping Li, Ph.D., Warren V. Kalina, Ph.D., David Cooper, Ph.D., Robert W. Frenck, Jr., M.D., Laura L. Hammitt, M.D., Özlem Türeci, M.D., Haylene Nell, M.D., Axel Schaefer, M.D., Serhat Ünal, M.D., Dina B. Tresnan, D.V.M., Ph.D., Susan Mather, M.D., Philip R. Dormitzer, M.D., Ph.D., Uğur Şahin, M.D., Kathrin U. Jansen, Ph.D., and William C. Gruber, M.D., for the C4591001 Clinical Trial Group*

Alternative history

- Imagine that BNT162b2 was developed and licensed before ChAdOx1
- We conduct an active-control trial to evaluate ChAdOx1, with BNT162b2 as the control arm
- Primary endpoint of COVID-19

Results from hypothetical trial

	BNT162b2	ChAdOx1
PYFU per arm	10,000	10,000
Observed COVID-19 cases	20	80
Rate ratio (90% CI)	REF	4.00 (2.61-6.32)

Results from hypothetical trial

	BNT162b2	ChAdOx1
PYFU per arm	10,000	10,000
Observed COVID-19 cases	20	80
Rate ratio (90% CI)	REF	4.00 (2.61-6.32)
Vaccine efficacy (%)	95	80
COVID-19 cases if subjects had not been vaccinated	400	400
COVID-19 cases averted	380	320

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 MARCH 14, 2019 VOL. 380 NO. 11

One Month of Rifapentine plus Isoniazid to Prevent HIV-Related Tuberculosis

S. Swindells, R. Ramchandani, A. Gupta, C.A. Benson, J. Leon-Cruz, N. Mwelase, M.A. Jean Juste, J.R. Lama, J. Valencia, A. Omoz-Oarhe, K. Supparatpinyo, G. Masheto, L. Mohapi, R.O. da Silva Escada, S. Mawlana, P. Banda, P. Severe, J. Hakim, C. Kanyama, D. Langat, L. Moran, J. Andersen, C.V. Fletcher, E. Nuermberger, and R.E. Chaisson, for the BRIEF TB/A5279 Study Team*

BRIEF trial – summary

- Compared short regimen (1-month) against standard (9-month) regimen
- Primary endpoint: diagnosis of tuberculosis, or death from tuberculosis or unknown cause
- NI margin: 1.25 per 100 PYFU for the rate difference
- Assumed incidence of 2 per 100 PYFU in the 9-month arm

Group	PFYU	Endpoints	Incidence rate (per 100 PYFU)	Rate difference (95% CI)
9-month	4896	33	0.67	REF
1-month	4926	32	0.65	-0.02 (-0.35, 0.30)

Interpretation

- “We found that 1 month of daily rifapentine plus isoniazid was noninferior to daily isoniazid for 9 months for the prevention of tuberculosis in HIV-infected adults and adolescents.”
- Results are consistent with two explanations
  - the two regimens are both effective (to a similar degree)
  - both regimens are ineffective

HIV pre-exposure prophylaxis (PrEP)

- Taking antiretroviral drugs around the time of sex to prevent the acquisition of infection
- TDF-FTC (Truvada) was first drug approved for PrEP
- Strong evidence that TDF-FTC is **highly** protective (>95%) if adhere to regimen
- Several active-control trials have been performed (with TDF-FTC as control arm)

Study	Incidence rate*		Rate ratio (90% CI)	
	TDF-FTC	PrEP X		
A	0 (0)	0 (0)	Undefined	
B	1 (20)	1 (20)	1.00 (0.56,1.77)	
C	2 (40)	2 (40)	1.00 (0.66,1.48)	

* per 100 PY (number of endpoints)
2000 PYFU per arm

Study	Incidence rate*			Rate ratio (90% CI)	
	Placebo	TDF-FTC	PrEP X		
A	5	0 (0)	0 (0)	Undefined	
B	5	1 (20)	1 (20)	1.00 (0.56,1.77)	
C	5	2 (40)	2 (40)	1.00 (0.66,1.48)	

* per 100 PY (number of endpoints)
2000 PYFU per arm

Study	Incidence rate*			Rate ratio (90% CI)	
	Placebo	TDF-FTC	PrEP X		
A		1 (20)	2 (40)	2.00 (1.27,3.14)	
B		1 (20)	2 (40)	2.00 (1.27,3.14)	
C		1 (20)	2 (40)	2.00 (1.27,3.14)	

* per 100 PY (number of endpoints)

2000 PY observation per arm

Study	Incidence rate*			Rate ratio (90% CI)	
	Placebo	TDF-FTC	PrEP X		
A	2	1 (20)	2 (40)	2.00 (1.27,3.14)	
B	5	1 (20)	2 (40)	2.00 (1.27,3.14)	
C	10	1 (20)	2 (40)	2.00 (1.27,3.14)	

* per 100 PY (number of endpoints)
2000 PY observation per arm

Conclusion

- Valid interpretation of an active-control HIV PrEP trial must consider the HIV incidence that **would have been observed** in a hypothetical placebo group (counterfactual HIV incidence)

Averted infections ratio (AIR)

$$\mathrm{AIR} = \frac{\lambda_P - \lambda_E}{\lambda_P - \lambda_C}$$

λ = incidence rate

P= placebo arm, E= experimental arm, C = control arm (TDF-FTC)

- AIR measures the proportion of infections that would be averted by using the experimental drug rather than the control drug

  AIR = 1: two drugs equally effective

  AIR < 1: new drug less effective

  AIR > 1: new drug more effective

- Natural "preservation of effect" measure for non-inferiority trials
- Conclusions about non-inferiority based on the lower confidence limit

Study	Incidence rate*		Rate ratio (90% CI)	
		TDF-FTC	PrEP X	
A	5	0 (0)	0 (0)	Undefined	
B	5	1 (20)	1 (20)	1.00 (0.56,1.77)	
C	5	2 (40)	2 (40)	1.00 (0.66,1.48)	

* per 100 PY (number of endpoints)
2000 PYFU per arm

Study	Incidence rate*			Rate ratio (90% CI)	AIR (90% CI)
	Placebo	TDF-FTC	PrEP X		
A	5	0 (0)	0 (0)	Undefined	1.00 (0.98,1.02)
B	5	1 (20)	1 (20)	1.00 (0.56,1.77)	1.00 (0.88,1.14)
C	5	2 (40)	2 (40)	1.00 (0.66,1.48)	1.00 (0.78, 1.28)

* per 100 PY (number of endpoints)
2000 PY observation per arm

Study	Incidence rate*		Rate ratio (90% CI)	
		TDF-FTC	PrEP X	
A	2	0 (0)	0 (0)	Undefined	
B	5	1 (20)	1 (20)	1.00 (0.56,1.77)	
C	10	2 (40)	2 (40)	1.00 (0.66,1.48)	

* per 100 PY (number of endpoints)
2000 PYFU per arm

Study	Incidence rate*			Rate ratio (90% CI)	AIR (90% CI)
	Placebo	TDF-FTC	PrEP X		
A	2	1 (20)	2 (40)	2.00 (1.27,3.14)	0.0 (ND)
B	5	1 (20)	2 (40)	2.00 (1.27,3.14)	0.75 (0.62,0.91)
C	10	1 (20)	2 (40)	2.00 (1.27,3.14)	0.89 (0.82,0.96)

* per 100 PY (number of endpoints)
2000 PY observation per arm

Emtricitabine and tenofovir alafenamide *vs* emtricitabine and tenofovir disoproxil fumarate for HIV pre-exposure prophylaxis (DISCOVER): primary results from a randomised, double-blind, multicentre, active-controlled, phase 3, non-inferiority trial

Kenneth H Mayer, Jean-Michel Molina, Melanie A Thompson, Peter L Anderson, Karam C Mounzer, Joss J De Wet, Edwin DeJesus, Heiko Jessen, Robert M Grant, Peter J Ruane, Pamela Wong, Ramin Ebrahimi, Lijie Zhong, Anita Mathias, Christian Callebaut, Sean E Collins, Moupali Das, Scott McCallister, Diana M Brainard, Cynthia Brinson, Amanda Clarke, Pep Coll, Frank A Post, C Bradley Hare

Summary

Background Tenofovir alafenamide shows high antiviral efficacy and improved renal and bone safety compared with tenofovir disoproxil fumarate when used for HIV treatment. Here, we report primary results from a blinded phase 3 study evaluating the efficacy and safety of pre-exposure prophylaxis (PrEP) with emtricitabine and tenofovir alafenamide versus emtricitabine and tenofovir disoproxil fumarate for HIV prevention.

Lancet 2020; 396: 239–54

See Comment page 214

The Fenway Institute, Fenway Health, Boston, MA, USA (Prof K H Mayer MD);

DISCOVER – design

- Meta-analysis of three previous HIV prevention studies of TDF/FTC versus placebo, yielded
  - expected HIV incidence of 1.44 infections per 100 PYFU in TDF/FTC group
  - rate ratio between the placebo and TDF/FTC groups of 5.1 (95% CI 2.64–9.70).
- Non-inferiority margin of 1.62 to preserve at least 50% of the effect of TDF/FTC
- 5000 PFYU per arm achieves 82.5% power to establish non-inferiority of TAF/FTF to TDF/FTC

DISCOVER – primary results

Group	No. subjects	PYFU	Incident HIV infections	Incidence rate (per 100 PYFU)	Rate ratio (95% CI)
TDF/FTC	2693	4386	11	0.251	REF
TAF/FTC	2694	4370	6	0.137	0.55 (0.20,1.48)

- Incidence much lower than expected
- Established non-inferiority (1.48 < 1.62) <u>just</u>

DISCOVER – tweaked results

	Incident HIV infections		Rate ratio (95% CI)	Inference
	TDF/FTC	TAF/FTC		
Observed data	11	6	0.55 (0.20,1.48)	Non-inferiority of TAF/FTC demonstrated
1 additional infection*	11	7	0.64 (0.25,1.65)	Non-inferiority of TAF/FTC not demonstrated
3 fewer infections*	11	3	0.27 (0.08,0.98)	Superiority of TAF/FTC demonstrated

* in TAF/FTC arm

A Bayesian averted infection framework for PrEP trials with low numbers of HIV infections: application to the results of the DISCOVER trial

David V Glidden, Oliver T Stirrup, David T Dunn

Trials of candidate agents for HIV pre-exposure prophylaxis (PrEP) might randomly assign participants to be given a new PrEP agent or oral coformulated tenofovir disoproxil fumarate plus emtricitabine. This design presents unique challenges in interpretation. First, with two active arms, HIV incidence might be low. Second, the effectiveness of tenofovir disoproxil fumarate plus emtricitabine varies across populations; thus, similar HIV incidence between groups could be consistent with a wide range of effectiveness for the new PrEP. We propose a two-part approach to trial results. First, we use Bayesian methods to incorporate assumptions about the background incidence of HIV in the trial in the absence of PrEP, possibly augmented by external data. On the basis of the estimated background incidence, we estimate and compare the number of averted (or prevented) HIV infections in each of the two trial groups, calculating the averted infections ratio. We apply these methods to a completed trial of tenofovir alafenamide plus emtricitabine for PrEP. Our framework shows that leveraging external information to estimate averted infections and the averted infections ratio enhances the efficiency and interpretation of active-controlled PrEP trials.

Lancet HIV 2020; 7: e791–96

Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA (Prof D V Glidden PhD) **and Institute for Global Health** (O T Stirrup PhD) **and MRC Clinical Trials Unit** (Prof D T Dunn PhD), **University College London, London, UK**

Correspondence to:
Prof David V Glidden,
Department of Epidemiology

Averted infections ratio (AIR)

$$\mathrm{AIR} = \frac{\lambda_P - \lambda_E}{\lambda_P - \lambda_C}$$

λ = incidence rate

P= placebo arm, E= experimental arm, C = control arm (TDF-FTC)

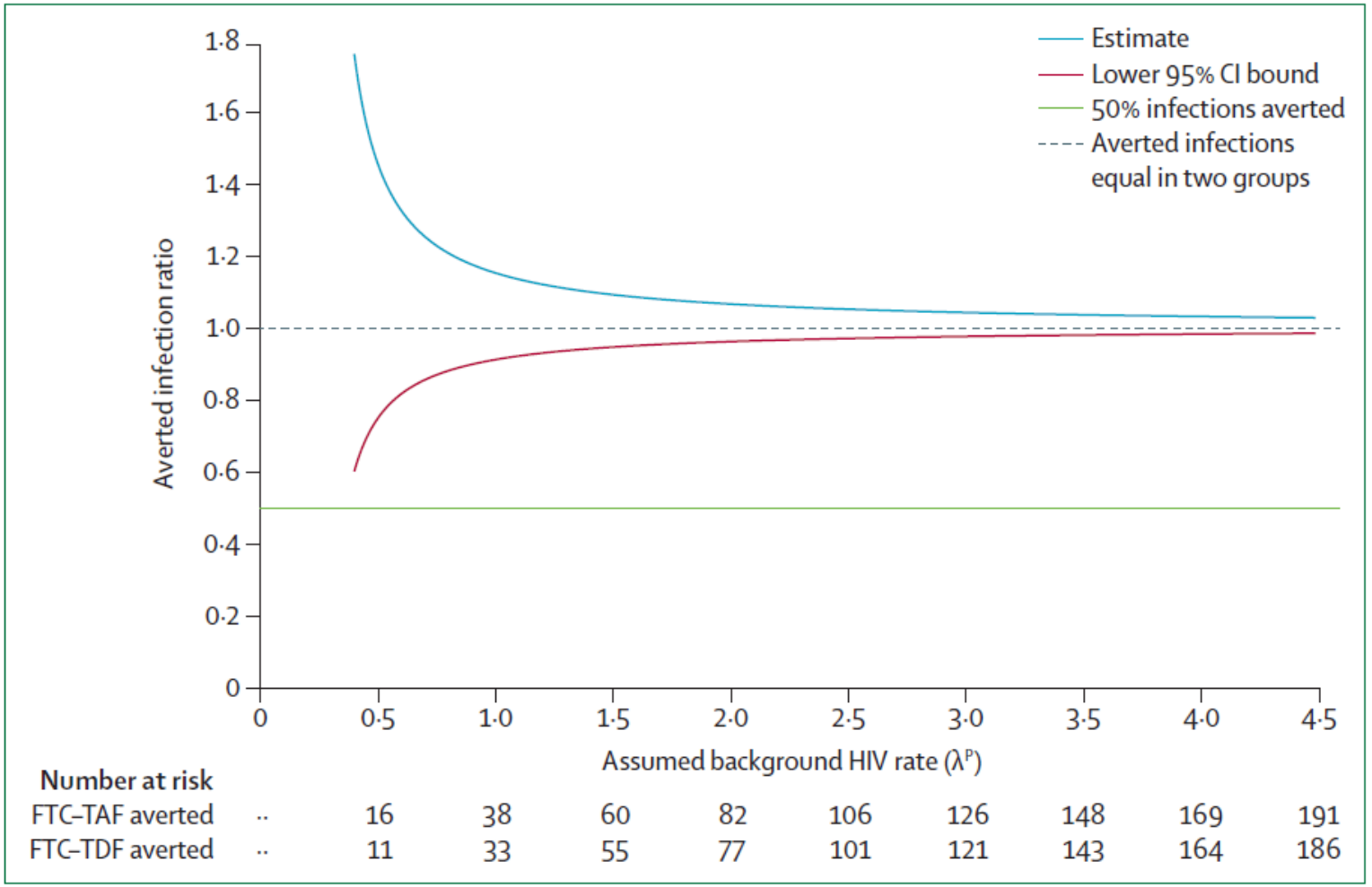


Figure 2: **Effect of varying assumed background HIV incidence on interpretation of DISCOVER data**
Averted infections ratio and lower 95% confidence limit were calculated from the on-study infections. The horizontal line at 0·5 demarcates the region for which tenofovir alafenamide plus emtricitabine averted at least 50% of infections—a measure of 50% effect preservation and thus evidence for non-inferiority. FTC–TAF=tenofovir alafenamide fumarate plus emtricitabine. FTC–TDF=tenofovir disoproxil fumarate plus emtricitabine.

DISCOVER – tweaked results

	Incident HIV infections		Rate ratio (95% CI)	Inference
	TDF/FTC	TAF/FTC		
Observed data	11	6	0.55 (0.20,1.48)	Non-inferiority of TAF/FTC demonstrated
1 additional infection*	11	7	0.64 (0.25,1.65)	Non-inferiority of TAF/FTC not demonstrated
3 fewer infections*	11	3	0.27 (0.08,0.98)	Superiority of TAF/FTC demonstrated

* in TAF/FTC arm

DISCOVER – tweaked results

	Incident HIV infections		Rate ratio (95% CI)	AIR (95% CI)
	TDF/FTC	TAF/FTC		
Observed data	11	6	0.55 (0.20,1.48)	1.10 (0.94,1.17)
1 additional infection*	11	7	0.64 (0.25,1.65)	1.08 (0.92,1.26)
3 fewer infections*	11	3	0.27 (0.08,0.98)	1.15 (1.02, 1.31)

* in TAF/FTC arm

Estimating counterfactual incidence

Approach	Examples
Epidemiological surveillance or data from prior prospective study in a similar population	Baeten et al. PLOS Med 2016.
Registrational cohort from which patients are randomised	PrEPVacc trial
Recency assay applied to baseline samples	Gao et al. Stat Comm Infect Dis (in press)
Ecological association between incidence of HIV and other STIs	Mullick et al. J Inf Dis 2019.
Association between adherence to TDF/FTC and preventative efficacy	Glidden et al. J Int AIDS Soc 2021.

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 MARCH 14, 2019 VOL. 380 NO. 11

One Month of Rifapentine plus Isoniazid to Prevent HIV-Related Tuberculosis

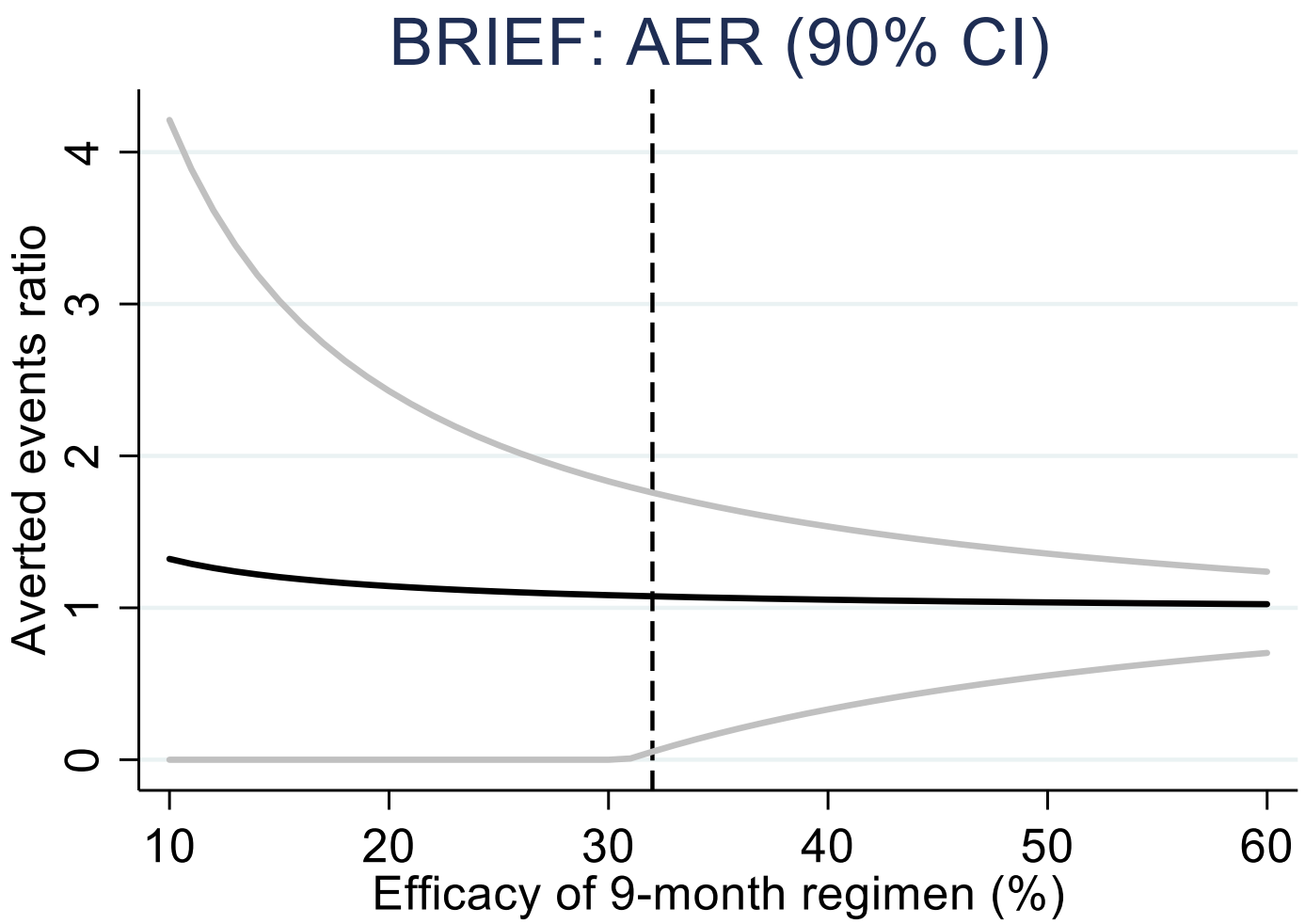
S. Swindells, R. Ramchandani, A. Gupta, C.A. Benson, J. Leon-Cruz, N. Mwelase, M.A. Jean Juste, J.R. Lama, J. Valencia, A. Omoz-Oarhe, K. Supparatpinyo, G. Masheto, L. Mohapi, R.O. da Silva Escada, S. Mawlana, P. Banda, P. Severe, J. Hakim, C. Kanyama, D. Langat, L. Moran, J. Andersen, C.V. Fletcher, E. Nuermberger, and R.E. Chaisson, for the BRIEF TB/A5279 Study Team*

Alternative formula for AIR

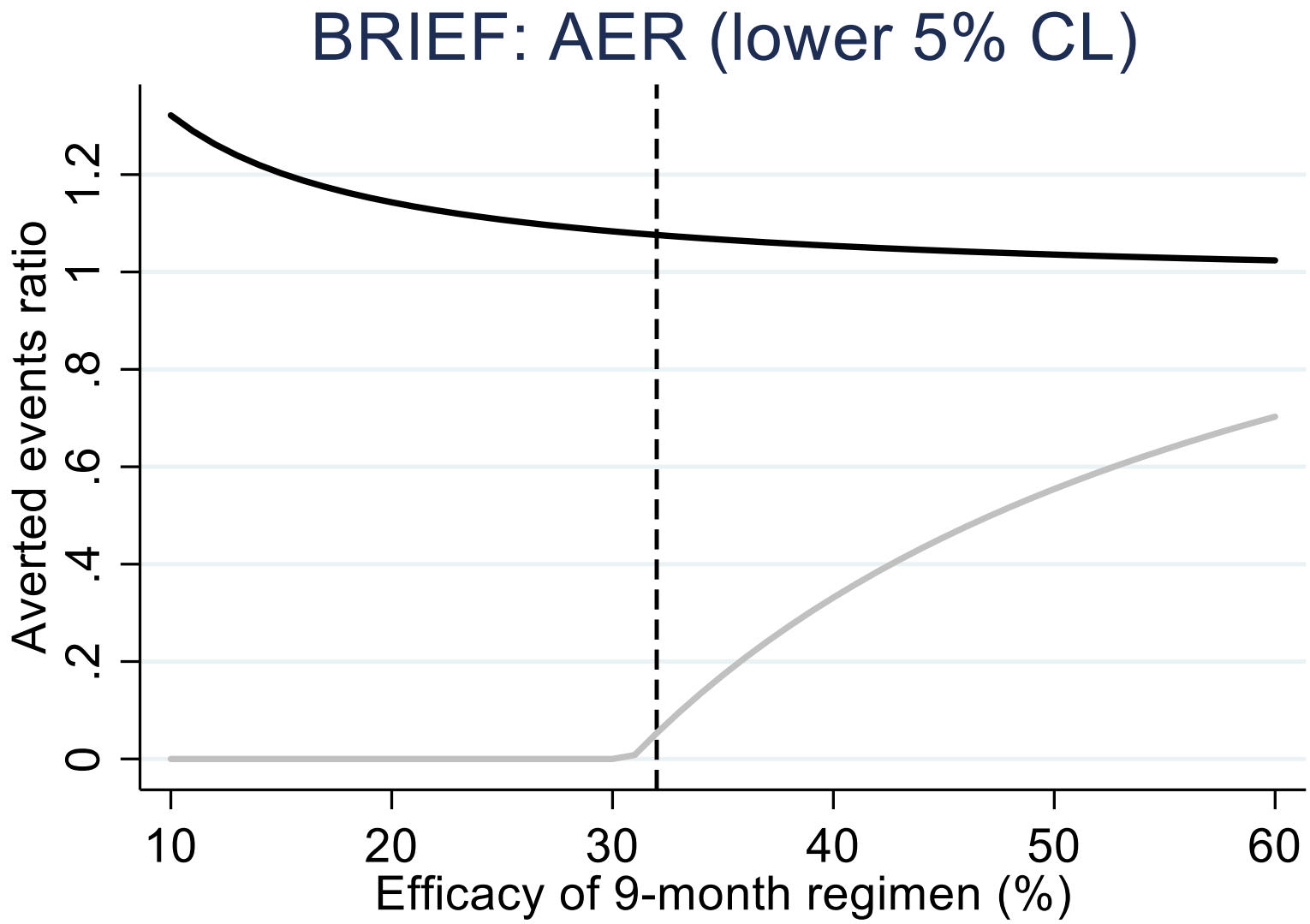
θ_C = (counterfactual) effectiveness of the active control drug in the **current** trial

$$\theta_C = 1 - \lambda_C/\lambda_P$$

$$\text{AIR} = \frac{\lambda_C - \lambda_E + \theta_C \lambda_E}{\theta_C \lambda_C}$$



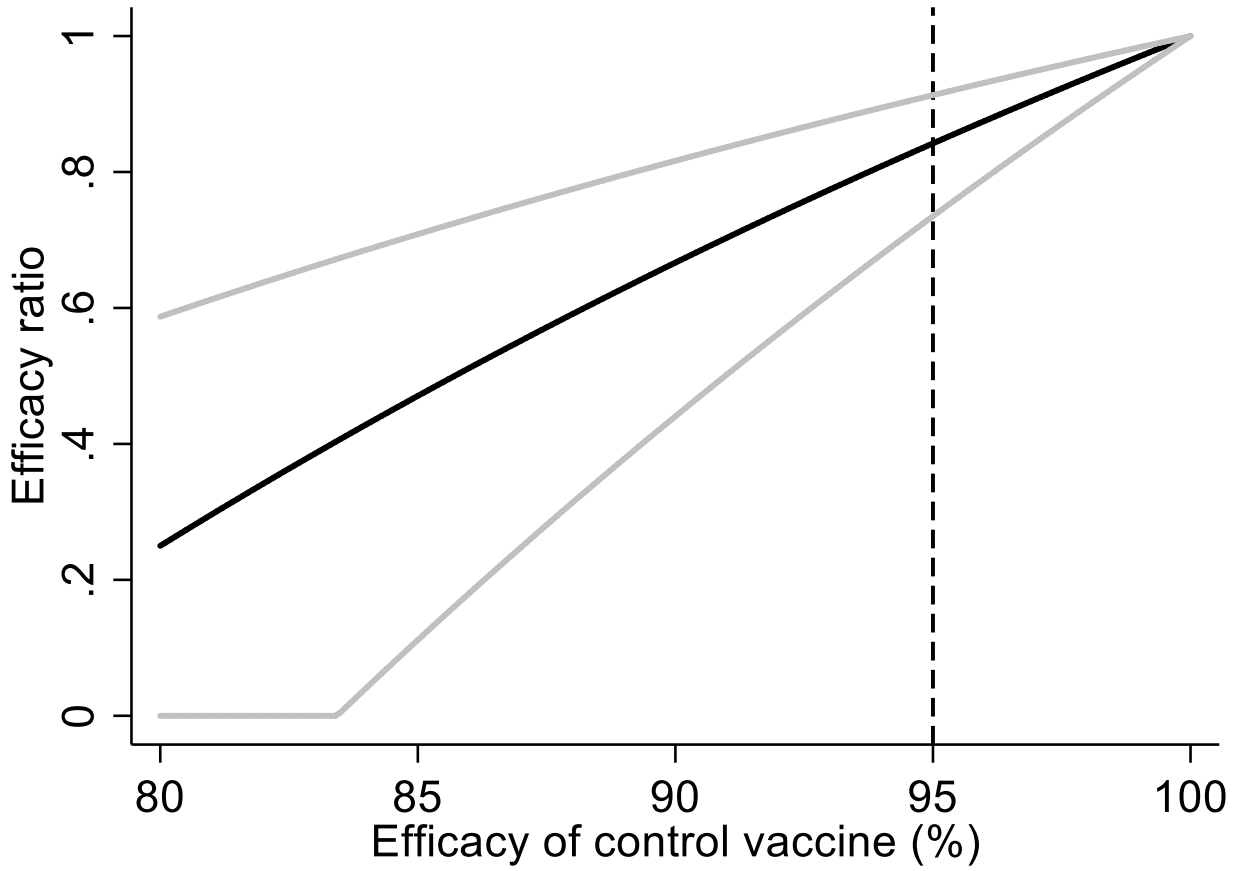
Meta-analysis reported efficacy of 32% (95% CI 5-51%). Ross et al. Lancet HIV 2021; 8: e8–15


BRIEF: AER (lower 5% CL)
Averted events ratio
0
.2
.4
.6
.8
1
1.2
10
20
30
40
50
60
Efficacy of 9-month regimen (%)

Results from hypothetical trial

	BNT162b2	ChAdOx1
PYFU per arm	10,000	10,000
Observed COVID-19 cases	20	80
Rate ratio (90% CI)	REF	4.00 (2.61-6.32)

Averted events ratio (90% CI) : ChAdOx1 versus BNT162b2



Alternative interpretation

$$\text{AIR} = \frac{\text{Efficacy of experimental treatment}}{\text{Efficacy of control treatment}}$$

Counterfactual efficacy

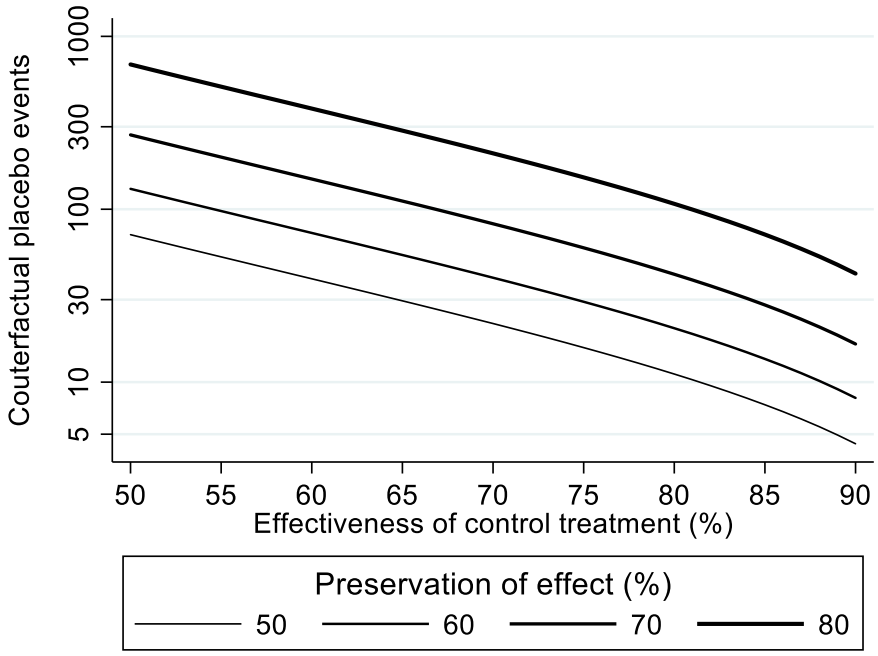
- Require estimate of efficacy in the current active-control trial
- Cannot naively extrapolate efficacy estimates from earlier placebo-controlled trials
- "Constancy" assumption
- For COVID-19 vaccines, issues include:
  - different circulating strains of SARS-CoV-2
  - trial population characteristics related to immune response
  - definition of clinical endpoints
  - evolution in disease symptomatology
  - interval between vaccination and viral exposure

Factors affecting sample size using AIR

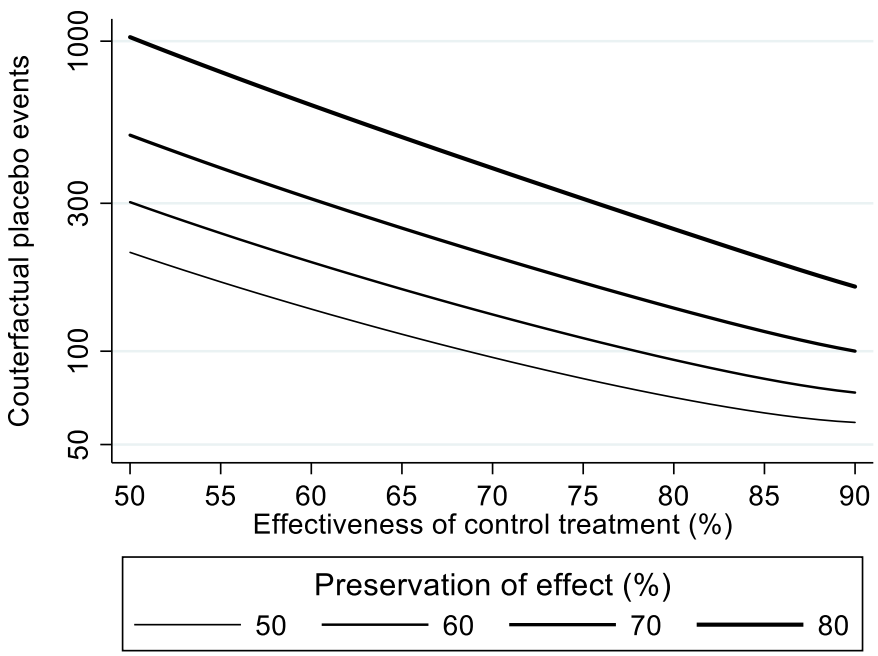
- Counterfactual placebo incidence
- Efficacy of control treatment
- “Preservation of effect” size
- Lower bound of CI for asserting non-inferiority
- Statistical power
- Whether estimating AIR via counterfactual placebo incidence or counterfactual efficacy

Sample size based on:

AIR estimated via placebo incidence

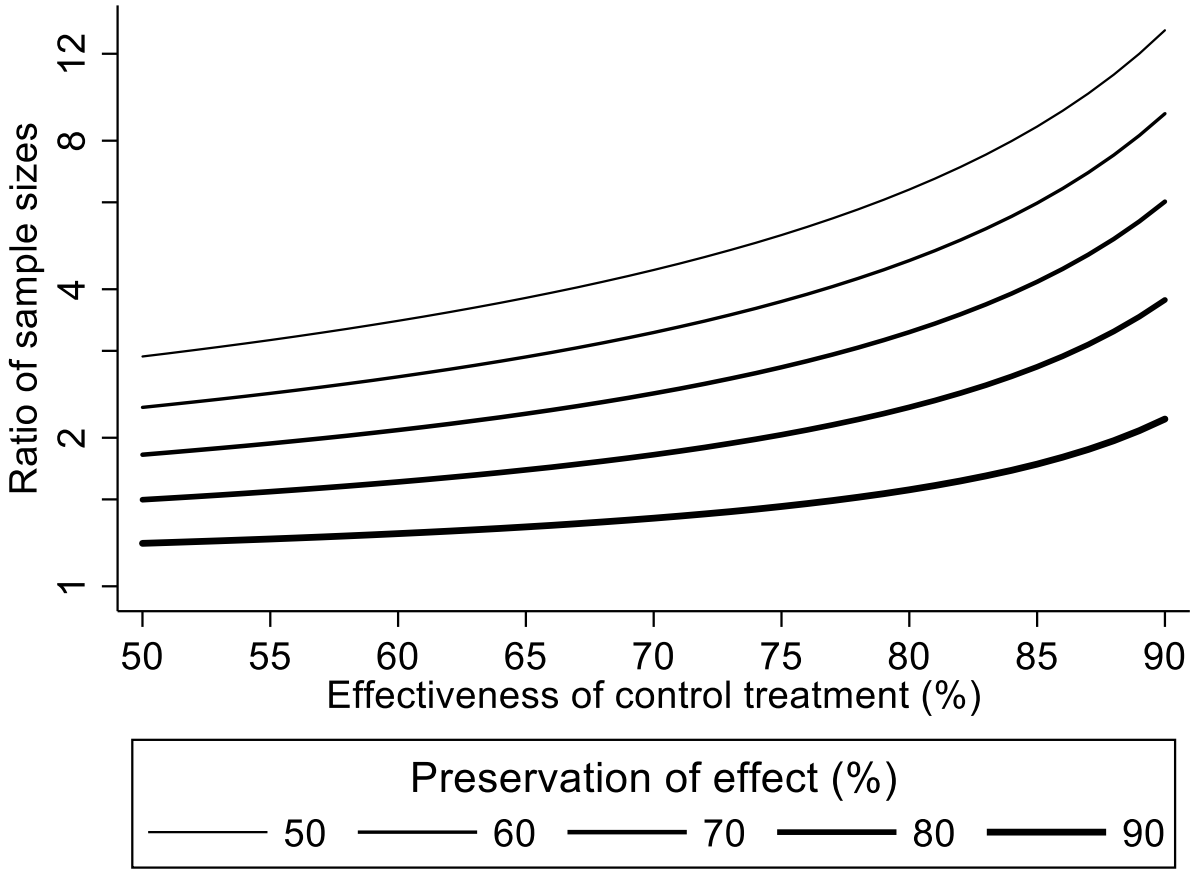


AIR estimated via control arm efficacy



NI based on lower 5% CL, 90% power

Ratio of sample sizes: AIR estimated via counterfactual efficacy versus placebo incidence



Conclusions

- Standard analytical approach for active-control trials with time-to-event endpoint can be clinically misleading
- Important to use a metric that includes either counterfactual placebo incidence or control treatment efficacy
- Using counterfactual placebo incidence much more powerful
- Logic presumably extends to other types of endpoint
- Do problems exist where the standard approach is actually valid?

Acknowledgements

David Glidden (UCSF)
Sheena McCormack (UCL)
Oliver Stirrup (UCL)

References

Dunn DT, Glidden DV, Stirrup OT, McCormack S. The averted infections ratio: a novel measure of effectiveness of experimental HIV pre-exposure prophylaxis agents. Lancet HIV. 2018;5(6):e329-e34.

Dunn DT, Glidden DV. The Connection between the Averted Infections Ratio and the Rate Ratio in Active-control Trials of Pre-exposure Prophylaxis Agents. Statistical Communications in Infectious Diseases. 2019;11(1) 20190006

Glidden DV, Stirrup OT, Dunn DT. A Bayesian averted infection framework for PrEP trials with low numbers of HIV infections: application to the results of the DISCOVER trial. Lancet HIV. 2020;7(11):e791-e6.

Dunn DT, Stirrup OT, Glidden DV. Confidence limits for the averted infections ratio estimated via the counterfactual placebo incidence rate. Statistical Communications in Infectious Diseases 2021; 13(1): 20210002

Glidden DV, et al. Using the adherence-efficacy relationship of emtricitabine and tenofovir disoproxil fumarate to calculate background HIV incidence: a secondary analysis of a randomized, controlled trial. J International AIDS Society 2021, 24:e25744.