Published in final edited form as:
AIDS. 2012 March 13; 26(5): 529–532. doi:10.1097/QAD.0b013e3283509a29.

Design challenges facing clinical trials of the effectiveness of new HIV prevention technologies

Anneke C. GROBLER, MSc[1] and **Salim S. ABDOOL KARIM, MB ChB, PhD**[1,2]

[1]Centre for the AIDS Programme of Research in South Africa (CAPRISA), University of KwaZulu-Natal, Durban, South Africa,

[2]Department of Epidemiology, Columbia University, New York, USA

Abstract

Recent successes of antiretroviral pre-exposure prophylaxis (PrEP) in preventing HIV infection have raised questions whether further placebo controlled trials of new HIV-prevention technologies are ethically justifiable. A trial with active agent(s) in the comparator group can be designed either as a superiority or non-inferiority trial. In a non-inferiority trial the hypothesis tested is that the intervention is not inferior to, by a predefined clinically relevant amount, or at least as effective as, the comparator. Non-inferiority trials pose challenges in data interpretation.

Firstly it is possible to show equivalence of two non-effective interventions. If the active comparator intervention is ineffective, the new intervention would be shown to be non-inferior to this inactive intervention, while neither intervention is superior to placebo or no intervention. The second challenge is that any effect that dilutes the true efficacy of an intervention in a trial, such as non-adherence, loss to follow-up or protocol violations, makes it easier for the two interventions to be declared equivalent. Non-differential low adherence is unlikely to lead to the conclusion that an inferior intervention is non-inferior. However, differential adherence between study arms, which is more likely in non-blinded trials, is likely to bias the results and lead to incorrect conclusions.

Investigators conducting non-inferiority trials will have to pay special attention to supporting, measuring and maintaining high adherence. The goal in future non-inferiority trials should be to maintain similar levels of high adherence in all study arms, but at a minimum to reduce the likelihood of differential adherence across study arms.

INTRODUCTION

Several randomized placebo-controlled trials of microbicides or antiretroviral pre-exposure prophylaxis (PrEP) have been conducted to assess their effectiveness in preventing sexual transmission of HIV infection. Prior to 2010, none of these trials had demonstrated effectiveness in preventing HIV [1]. In July 2010, a microbicide gel containing the antiretroviral drug tenofovir (CAPRISA 004) demonstrated 39% reduction in HIV-acquisition in women [2] and in November 2010, oral emtricitabine and tenofovir disoproxil fumarate (FTC–TDF) showed 44% reduction in HIV-acquisition (iPrEx) in men who have sex with men [3]. While oral FTC-TDF (FEM-PrEP) [4] and oral TDF (VOICE) [5] were not found to be effective in heterosexual women, two trials showed in July 2011 that oral TDF and FTC-TDF were found to reduce HIV-transmission by 62% and 73% in serodiscordant couples (Partners PrEP) [6] and oral FTC-TDF reduced HIV-transmission by

Name and address for correspondence: Anneke Grobler, CAPRISA, 2nd Floor DDMRI, Nelson R Mandela School of Medicine, University of KwaZulu-Natal, Private Bag X7, Congella, 4013. grobler@ukzn.ac.za.

63% in heterosexual men and women (TDF2) [7]. Furthermore, early antiretroviral treatment initiation in the HIV-positive partner (HPTN052) was shown to reduce HIV acquisition from the positive partner in the HIV-negative partner in serodiscordant couples by 96% [8]. These recent successes have raised questions about the design of future HIV prevention trials and whether further placebo controlled trials of some new HIV-prevention technologies are ethically justifiable [9].

As long as evidence about the efficacy of PrEP in preventing HIV infection is ambiguous, future placebo controlled trials are required. Currently estimates of the efficacy of daily oral FTC-TDF in women ranges from no efficacy [4] to 73% [6]. However, there is hope that accumulating evidence may demonstrate a more consistent estimate of PrEP effectiveness in the near future. In this scenario, future placebo controlled trials may no longer be possible.

One option for future trials which are designed to assess the efficacy of new HIV prevention strategies is to include these efficacious antiretroviral intervention(s) in both intervention and control arms. In this case, the net effect is to lower the overall HIV incidence rate in the trial thereby making the trial larger and/or longer. This could add substantially to the cost of the trial. In some instances, it may not be safe, feasible or practical to implement a known efficacious antiretroviral intervention simultaneously in a trial as it may interfere with the new study intervention under investigation. For example, the use of a vaginal ring containing an antiretroviral may be a contra-indication for the simultaneous use with a gel containing the same or a different antiretroviral. In such cases, another option is to include efficacious antiretroviral intervention(s) only in the comparator group.

A trial with active agent(s) in the comparator group can be designed either as a superiority or non-inferiority trial. In a superiority trial, the hypothesis tested is that the new intervention under investigation is better, by a clinically relevant amount, than the comparator, which can be either standard of care, placebo or another active intervention. This can be done either as a blinded or open label trial. The new intervention needs to be even more potent than an active comparator in order to show superiority i.e. higher efficacy.

In a non-inferiority trial the hypothesis being tested is that the intervention under investigation is not inferior to, by a predefined clinically relevant amount, or at least as effective as, the comparator [10]. Non-inferiority trials are used when the new intervention is assumed to have some benefits (e.g. improved safety, low cost, etc) over the comparator intervention, but has similar efficacy [11]. Non-inferiority trials, however, are not without challenges in data interpretation [12].

The first challenge is that any intervention can easily be shown to be non-inferior to a non-effective comparator intervention; whether or not either intervention is superior to placebo or no intervention [10].

The assumption made when an active comparator is used in a non-inferiority trial is that the efficacy found in another trial also applies to the population and environment in the current trial; an assumption that cannot, in most instances be readily tested. How similar do two populations need to be before findings in one population can be used to determine the active comparator in another population? For example, can we assume that oral FTC–TDF will be 44% effective in a population of heterosexual women, because it was found to be 44% effective in a population of men who have sex with men? [3] The validity of this assumption was challenged by the findings of a trial testing FTC-TDF in a population of heterosexual women (FEM-PrEP) where no protection against HIV was shown [4], while 73% effectiveness was found in men and women in serodiscordant relationships (Partners PrEP) [6].

Consider a hypothetical new PrEP trial testing coital use of oral tenofovir against daily use of oral tenofovir which concluded equivalence; i.e. concluded that coital use of oral tenofovir is not inferior to daily oral tenofovir. If we drew this conclusion in August 2011, before the recent findings of the VOICE study [5], we would have concluded that coital tenofovir is 62% effective, based on the effectiveness of daily oral tenofovir in the Partners PrEP trial [6]. If, however, we drew this conclusion 3 months later in October 2011, after the recent announcement from the VOICE study, we could conclude that coital tenofovir is not effective. Without a clear and consistent protective efficacy estimate for daily oral tenofovir, it is difficult to provide a meaningful interpretation of a non-inferiority trial assessing coital use of oral tenofovir.

The second challenge is that any effect that dilutes the true efficacy of an intervention in a trial such as non-adherence to the study regimen, considerable loss to follow-up or protocol violations, makes it easier for the two interventions to be declared equivalent. Non-adherence is a particular concern in these settings because it difficult to reliably measure adherence levels in both arms of a PrEP trial. Drug level monitoring is only a partial solution, since this can only be done in the active arm. A placebo with a biological marker of adherence could address this. A poorly conducted trial will be more likely to lead to the false conclusion that the test intervention is non-inferior to the active comparator than a well conducted trial [10].

The standard approach to dealing with this problem is to base the primary analysis of a non-inferiority trial on the per protocol population and not on the intent-to-treat population, as is standard with a superiority trial. This is because the intent-to-treat analysis is biased towards equality in conditions where many participants did not follow study procedures. One of the most important criteria for inclusion in the per protocol analysis is adherence to the study regimen. In the CAPRISA 004 [2], iPrEx [3], Partners PrEP [6] and FEM-PrEP trials [4], self-reported adherence and adherence based on applicator or pill count was high. In the iPrEx trial there was a poor correlation between reported adherence and drug levels detected. This casts doubts on both the adherence during this trial and on the validity of self-report. If adherence is not accurately measured, an analysis based on adherence or on the per protocol population determined by adherence may lead to incorrect conclusions. More reliable measures of adherence such as drug concentrations in the active arm could go some way towards increasing the likelihood of high adherers being included in the active arm of the per protocol analysis. A biological adherence marker will be needed for the placebo arm for a valid comparison.

The importance of adherence in non-inferiority trials should not be underestimated. Sub-optimal adherence in a trial with an active comparator has the net effect of making the result “flatter” but does not lead to an erroneous conclusion of non-inferiority, provided adherence levels are similar in each of the study arms (Table 1). The exception occurs when adherence is zero or very close to zero in each of the treatment arms; in this case non-inferiority is invariably declared.

Differential adherence in the study arms is much more complex and, under a range of scenarios, can lead to incorrect conclusions (Table 1). If the true effectiveness of both interventions are the same, the intervention with higher adherence is favored in a non-inferiority study. If an inferior, but still efficacious, intervention had higher adherence than the more effective comparator, a conclusion of non-inferiority would be made.

Adherence to an intervention is an important determinant of whether a prevention strategy would have a public health benefit; therefore the adherence to an intervention is an important aspect of the determination of its efficacy. However, if adherence to the

comparator is low; one might in effect be comparing a new intervention to a placebo-like effect created by lack of adherence to the intervention even though this comparator was intended to be an active intervention.

Differential adherence is unlikely in a double blinded trial of similar interventions; where neither the study participants, nor the investigators are aware of study assignment. However, differential adherence is much more likely where the comparator and the new intervention are substantially different and blinding is less likely. More specifically, if one intervention is an oral formulation and the other intervention is topical, different patterns of use might be likely and adherence might differ substantially. The same holds for different dosing strategies; for example where once daily dosing is compared to coitally dependent dosing. Differential adherence is also likely when the interventions have different side effect profiles. Although no evidence of differential adherence was found in the CAPRISA 004 trial [2], differential adherence was reported in the iPrEx trial in some of the early visits, with lower adherence in the active arm [3], probably due to drug side effects.

Extreme caution should be used when non-inferiority studies are planned comparing different dosing strategies, which are likely to lead to different adherence levels. It might not be possible to compare formulations that are very different, such as oral, ring and gel formulations, as they are likely to be used differentially.

The third challenge with non-inferiority trials is their large size, much larger than superiority trials [12]. The PrEP trials which have recently announced results were all designed as superiority trials with sample sizes between 1000 and 5000 participants. The topical PrEP trial targeted 92 HIV infections[2], while the oral PrEP trial targeted 85 infections. In contrast, a non-inferiority trial with 80% power to show an intervention not more than 20% inferior to the active comparator would require about 500 HIV infections. A non-inferiority limit of 20% is probably too large if the active comparator is only 40% effective, but small enough if the active comparator is 60 to 70% effective.

Conclusions

In future efficacy trials of new HIV prevention interventions, non-inferiority designs may become one of the standard approaches. These designs are harder to interpret and care should be taken to ensure that comparison treatments are well understood.

In these study designs, adherence is a critical factor as it may lead to spurious results. Differential adherence in the treatment arms could lead to incorrect findings about the true effectiveness of the interventions. Investigators conducting non-inferiority trials will, of necessity, have to pay special attention to supporting, measuring and maintaining high adherence. The effect of differential adherence between study arms should be considered when interpreting the results.

AG and SSAK conceptualised this work. AG wrote the manuscript and SSAK provided comments and edits.

Acknowledgments

Anneke Grobler received training support from the Columbia University-Southern African Fogarty AIDS International Training and Research Programme (AITRP), funded by the Fogarty International Center (grant D43TW00231).

FUNDING The CAPRISA 004 Tenofovir gel trial was supported by the United States Agency for International Development (USAID), FHI (cooperative agreement GPO-A-00-05-00022-00 and contract 132119), CONRAD and LIFElab.

REFERENCES

1. Padian NS, McCoy SI, Balkus JE, Wasserheit JN. Weighing the gold in the gold standard: challenges in HIV prevention research. AIDS. 2010; 24:621–635. 610.1097/QAD. 1090b1013e328337798a. [PubMed: 20179575]
2. Abdool Karim Q, Abdool Karim SS, Frohlich JA, Grobler AC, Baxter C, Mansoor LE, et al. Effectiveness and Safety of Tenofovir Gel, an Antiretroviral Microbicide, for the Prevention of HIV Infection in Women. Science. 2010; 329:1168–1174. [PubMed: 20643915]
3. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, et al. Preexposure Chemoprophylaxis for HIV Prevention in Men Who Have Sex with Men. New England Journal of Medicine. 2010
4. FHI. FHI to Initiate Orderly Closure of FEM-PrEP. 2011. http://www.fhi.org/en/Research/Projects/FEM-PrEP.htm
5. National Institute of Allergy and Infectious Diseases. Statement: NIH Modifies 'VOICE' HIV Prevention Study in Women Oral Tenofovir Discontinued in Clinical Trial. 2011. http://www.nih.gov/news/health/sep2011/niaid-28.htm
6. University Of Washington International Clinical Research Center. Press release: Partners PrEP Study. Pivotal study finds that HIV medications are highly effective as prophyalxis against HIV infection in men and women in Africa. 2011. http://depts.washington.edu/uwicrc/research/studies/files/PrEP_PressRelease-UW_13Jul2011.pdf
7. Centers for Diseases Control and Prevention. Press release: CDC trial and another major study find PrEP can reduce risk of HIV infection among heterosexuals. Jul 13. 2011 http://www.cdc.gov/nchhstp/newsroom/PrEPHeterosexuals.html2011
8. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Prevention of HIV-1 Infection with Early Antiretroviral Therapy. New England Journal of Medicine. 2011
9. Kuhn L, Susser I, Stein Z. Can further placebo-controlled trials of antiretroviral drugs to prevent sexual transmission of HIV be justified? The Lancet. 2011; 378:285–287.
10. Schall, R.; Luus, H.; Erasmus, T. Type of comparison. In: Karlberg, J.; Tsang, K., editors. Introduction to clinical trials. The clinical trials centre; Hong Kong: 1998.
11. Lagakos, S.; Gable, A. Methodological challenges in biomedical HIV prevention trials. Institute of medicine of the national academies; Washington, DC: 2008.
12. Rothman M, Li N, Chen G, Chi G, Temple R, Tsou H. Design and analysis of non-inferiority mortality trials in oncology. Statistics in medicine. 2003; 22:239–264. [PubMed: 12520560]

AIDS. Author manuscript; available in PMC 2013 March 13.

Table 1

Projected effectiveness found with different levels of true efficacy of the comparator and new intervention in combination with different adherence levels. The study continued until 500 infections were observed in both arms combined.

Comparator intervention			New intervention				
Efficacy	**Adherence**	**Number of infections**	**Efficacy**	**Adherence**	**Number of infections**	**Effectiveness (95% CI)**	**Conclusion about new intervention**
Adherence similar							
50%	100%	250	50%	100%	250	0 (−19%; 16%)	Equally effective
	50%	250		50%	250	0 (−19%; 16%)	Equally effective
	25%	250		25%	250	0 (−19%; 16%)	Equally effective
	0%	250		0%	250	0 (−19%; 16%)	Equally effective
50%	100%	167	0%	100%	333	−99 (−140;−66)	Inferior
	50%	214		50%	286	−34(−60%-12%)	No more effective than control *
	25%	233		25%	267	−15 (−37; 4)	No more effective than control *
	0%	250		0%	250	0 (−19%; 16%)	Equally effective *
40%	100%	300	60%	100%	200	33 (20%; 44%)	Superior
	50%	266		50%	234	12 (−5%; 26%)	At least as effective as control
	25%	257		25%	243	5 (−13%; 21)	At least as effective as control
	0%	250		0%	250	0 (−19%; 16%)	Equally effective
40%	100%	215	20%	100%	285	−33 (−58%;−11%)	No more effective than control
	50%	235		50%	265	−13 (−34%; 5%)	No more effective than control
	25%	243		25%	257	−6 (−26%; 11)	No more effective than control
	0%	250		0%	250	0 (−19%; 16%)	Equally effective *
Adherence differential							
50%	50%	300	50%	100%	200	33 (20%; 44%)	Superior *
	100%	200		50%	300	−50 (−79%; −25%)	Inferior *
40%	50%	242	20%	100%	258	−7 (−27%; 11%)	No more effective than control
	25%	243		50%	257	−6 (−26; 11)	No more effective than control
	10%	267		80%	233	13 (−4; 27)	At least as effective as control *

Comparator intervention			New intervention				
Efficacy	**Adherence**	**Number of infections**	**Efficacy**	**Adherence**	**Number of infections**	**Effectiveness (95% CI)**	**Conclusion about new intervention**
70%	50%	241	30%	100%	259	−7 (−28; 10)	No more effective than control
	25%	246		50%	254	−3 (−23; 13)	No more effective than control
	10%	275		80%	225	18 (2; 31)	At least as effective as control *

Non-inferiority boundary of 20% used;

* incorrect decision made