

## **A Replication Plan for**

### **“Universal test and treat and the HIV epidemic in rural South Africa: a phase 4, open-label, community cluster randomized trial”**

*The Lancet*, HIV 5.3, e115-e126. DOI: [http://dx.doi.org/10.1016/S2352-3018\(17\)30205-9](http://dx.doi.org/10.1016/S2352-3018(17)30205-9)

By Collins C Iwuji MRCP, Joanna Orne-Gilemann PhD, Joseph Larmarange PhD, Eric Balestre MPH, Prof Rodolphe Thiebault PhD, Prof Frank Tanser PhD, Nonhlanhla PhD, Thembisa Makowa BA, Jaco Dryer NDip, Kobus Herbst Msc, Prof Nuala McGrath ScD, Prog Till Barninghausen, Sylvie Boyer PhD, Prof Tulio De Oliveria PhD, Claire Rekacewicz MD, Brigitte Bazin MD, Prof Marie-Louise Newell PhD, Prof Deenan Pillay PhD, Prof Francois Dabis PhD

Sridevi K. Prasad, MPH  
International Initiative for Impact Evaluation (3ie)  
Morgan Holmes, MSc  
International Initiative for Impact Evaluation (3ie)

August 2019

#### **Acknowledgements:**

We thank Dr. Anna C. Heard and Dr. John F. Creamer for their significant contributions in drafting this replication plan.

## 1. Introduction

In 2015, the World Health Organization (WHO) formally recommended global antiretroviral therapy (ART) for HIV-positive individuals as soon as they test positive. This recommendation calls for ART to be provided to all HIV-positive individuals regardless of CD4 count. Several trials in smaller settings, most notably Cohen et al. (2011), support the idea of widespread ART distribution as a means to effectively treat the HIV-affected population and reduce viral transmission rates.

The ANRS 12249 Treatment as Prevention trial (TasP) conducted by Iwuji et al. (2018) aimed to determine whether a test-and-treat program would be effective at reducing HIV incidence at the population level. The study, conducted in rural South Africa, was the first of four trials of its kind to report results (Havlir et al., 2019; Hayes et al., 2014; Moore et al., 2013). It aimed to treat the population by providing ART to all in a randomized setting, where HIV-positive individuals received ART no matter what their CD4 levels were. The control group received ART once their CD4 levels dropped to 350 cells/uL or less initially, and 500 cells/uL or less after January 2015 (Department of Health, 2014). The CD4 guideline changed after the results of the HPTN 052 and PARTNER studies were published, which showed a decrease in HIV incidence and transmission with early ART distribution between HIV-positive individuals and their serodiscordant partners (Cohen et al., 2011; Rodger et al., 2016).

The ANRS 12249 TasP finds a null effect on HIV incidence rates at the end of the six-year testing period in an area with an estimated 30% HIV prevalence rate. A deeper examination of the results showed poor linkage to care outcomes and high in- and out-migration, which likely contributed to the lack of clear program effects.

This replication study will use the raw data to reproduce the results in the original Iwuji et al. (2018) study. Additionally, it will apply different empirical methods to test the null result by examining the change in CD4 guidelines in 2015, the effects of the high migration in the study area, the effects of proximity to the nearest highway, and by using survival modelling techniques to look at time to HIV incidence. This plan continues with a further summary of Iwuji et al. (2018) in section 2. Section 3 presents the motivation for the replication and the methods we will employ in our study. Section 4 summarizes the work.

## 2. Presentation of the selected study

Iwuji et al. (2018) examine the use of Treatment as Prevention (TasP) for HIV-positive individuals in rural South Africa. The researchers contacted 26,518 participants (93% of eligible individuals) in 22 communities of KwaZulu-Natal, South Africa. Individuals were eligible to participate in the study if they spent four or more nights per week in one of the randomized clusters and were 16 years or older. Clusters were stratified by their estimated HIV prevalence rate and randomized to treatment or control within their HIV prevalence stratum. The study sites were local areas that encompassed many social and sexual networks. Additionally, the study observed in- and out-migration of the different communities, and collected information on sexual partners. The study took place in a six-

year period with four individual phases. The schedule follow-up lengths varied from two to four years depending on how early clusters were phased in (early cluster follow up was 4 years after baseline). All individuals in the study received access to counsellors at their point of care, rapid HIV counselling, and government-approved test kits at each round (mobile tests were introduced in the final survey).

The randomized component of the program was the delivery of ART for the treatment clusters, independent of their CD4 levels, in order to stem transmission to partners and potentially improve health in individuals with high CD4 counts. The control group received ART treatment based upon national guidelines. Pre-2015, this meant that initiation occurred once CD4 counts dropped to or below 350. The guidelines changed in January 2015, increasing the CD4 cut-off to 500. The treatment for these individuals began 2 weeks after identification unless they were seriously immunocompromised. Self-identified participants could continue on their normal course of treatment and all HIV-positive study participants were contacted by linkage to care teams if they did not attend a referred study clinic.

The main objective of the study is to understand how HIV incidence changed with the universal ART initiation becoming available at the population level. In addition, the study attempts to measure changes in HIV status ascertainment, linkage to care, and sexual behavioural changes. Modelling approaches simulated the necessary sample size to capture a 34% incidence rate reduction at 80% power (see Iwuji et al. [2018] for exact parameters). An intention-to-treat Poisson generalized estimating equation model estimates the marginal effect of the intervention on HIV incidence.

The results show that 93% of the selected individuals were contacted at least once, and were more likely to be women and older than average. 34% of these individuals out-migrated at some point during the study. Participants who out-migrated were more likely to be male and younger than average. 33% of the sample who were contacted were excluded from the incidence sample because of their first test sample was positive or the result was not valid. Of the remaining 67%, 80% had a follow-up test and were considered for the analysis of the incidence rate. This incidence sample was older than the median age and was more likely to be female than those who were not in the incidence group. The incidence rate in the sample was 2.2 (2.01-2.39), with an adjusted hazard ratio of 1.01 (0.87-1.17 95% confidence interval;  $p=0.89$ ). The crude mortality rate in the treatment group was 1.28 (0.84-1.72 95% confidence interval), where 33 deaths were reported, and 1.86 (1.38-2.34 95% confidence interval) in the control group, where 58 deaths were reported. There were 189 life-threatening or grade 4 clinic events, comprising 4% of each the treatment and control group.

The HIV care cascade in the study did not meet the UN standard, nor did the linkage of care reach the estimated 70% level (30% in each group). The authors suggest that the low linkage of care contributed to the lack of significant difference in treatment and control ART incidence. Additionally, the high rate of mobility (34%) could have made the null result more likely, as the care cascade struggles at the population level when in a smaller geographical area.

### **3. The proposed replication plan**

This study includes the standard objectives for 3ie-funded replication research (Brown, Cameron, & Wood, 2014). Our first step will be to complete a push-button replication to make sure that the author's code and data work as published in the study. Following this, we will conduct a pure replication to ensure that any changes in programming code and statistical software program do not affect the results. We then run a set of measurement and estimation analyses to assess if the results are consistent against the set of robustness checks described below. This will highlight any additional mechanisms beyond the poor linkage to care that may have contributed to the null result.

#### **3.1 Underlying rationale for the planned measurement and estimation analysis**

##### *3.1.1 Survival Analysis*

The original authors use an intention-to-treat Poisson generalized estimating equation modelling technique that takes cluster effects into account to assess the marginal effect of the treatment on HIV incidence. This provides a population-level estimate of the effect of the TasP treatment on HIV incidence by modelling the sum count of HIV seroconversions and total person-years. While the authors are able to incorporate cluster-level covariates, they do not use the individual-level data to see how time to HIV incidence is affected by the treatment group as they were looking at the population level. The authors use a GEE model that averaged the effects by cluster and therefore does not use individual data which is a weakness of the model chosen. We will use multilevel survival modelling techniques to take advantage of the individual-level and cluster-level data available. Survival analyses will allow us to look at the time to HIV incidence to see if treatment group had an effect on HIV transmission time.

##### *3.1.2 Change in ART initiation*

In January 2015, South Africa's Department of Health changed their HIV treatment guidelines to incorporate 2013 WHO guidelines recommending that ART be provided at CD4 counts under 500 cells/uL (Department of Health, 2014). In the primary manuscript and a separate commentary, the authors express their concerns on the effects that this guideline change may have on the effects of the TasP trial (Bärnighausen, Eyal, & Wikler, 2014; Iwuji et al., 2018). Since the implementation of this guideline would affect the control group, we will look at the HIV incidence rate changes before January 2015 and after January 2015 to see if the change in ART initiation contributed to the null result.

##### *3.1.3 Migration*

In their discussion, the authors highlight the high in- and out-migration rates in the study area as one potential driver of the null result. In other papers, the authors also identify high migration as the primary factor affecting improvements in the HIV care cascade (annual rates: out-migration 21.0%; in-migration 17.3%) (Larmarange et al., 2018). Additionally, in the original manuscript, those who out-migrated at least once were more likely to be younger, male, more educated, and actively seeking employment compared to those who never migrated (Table S4). Participants could migrate in and out of the study area multiple

times. They were still included in the incidence analysis and able to contribute person-time throughout the entire follow-up period. The dynamic population may have biased the results as they have poorer linkage to care and they may also have travelled to visit sexual partners outside of the study area. We will use methods to properly account for migration and see how migration affects the null result.

#### *3.1.4 Rural vs. Highway Area Incidence*

The authors note, in the discussion, the heterogeneity in prevalence rates between more rural areas and areas near highways. Tanser et al. (2009) show that HIV prevalence falls steeply as you move further away from main roads. The authors posit that policymakers should look to introduce TasP programs to areas with higher transmission rates to improve effectiveness, without presenting any results disaggregated by type of area. We plan to do this and examine whether the incidence rates vary in each of the study areas based upon this heterogeneity.

### **3.2 Methods**

#### *3.2.1 Pure Replication*

The pure replication will aim to reproduce the main tables (Tables 1-4; Table S7A) in the study. Replication of Table 1 will ensure that the sample is the same in the replication and original study, while Table 2 will report the replication incidence rate for HIV-positive tests by group and year. Table 3 will check the modelling assumptions, and Table 4 will estimate ART coverage for the trial. Replicating Table S7A will report the unadjusted and adjusted hazard ratios. The data was obtained from the Africa Health Research Institute data repository, and the code was provided by the authors. The code is for SAS, but was provided as a text file and translated for Stata use.

Any discrepancies between our work and the original authors will be resolved to the best of our ability through additional data work and communication with the original authors. If these discrepancies persist, we will note them in the report and comment on why they persist.

#### *3.2.2 Measurement and Estimation Analysis*

##### *3.2.2.1 Survival analysis*

We will use multilevel survival analysis to look at time to HIV incidence. We will first generate survival curves by treatment group, where participants will be right-censored if they do not develop HIV during the follow-up period. We will then test to see if the hazards are proportional. If they are, we will use a multivariable Cox regression to generate the hazard of developing HIV, controlling for individual-level and cluster-level covariates. If the proportional hazards assumption is violated, we will either use covariate-time interactions in the Cox regression model or use Accelerated Failure Time models. We will also check for the effects of unobserved heterogeneity and possible spurious causation in the model by including a frailty term in our multilevel Cox regression model (Austin, 2017).

### *3.2.2.2 Change in ART initiation*

Since the ART initiation policy was implemented starting in January 2015, we will split the incidence population into two groups at this cut-off point. First, we will compare linkage to care between the two groups. We will then generate incidence rates and hazard ratios for both time periods.

We will also generate survival curves by treatment group for both time periods to see if there is a difference in survival pre-2015 vs post-2015.

### *3.2.2.3 Migration*

The authors identify the difference in baseline characteristics between those who never out-migrated and those who migrated at least once (Table S4). Using the identified differences, we will identify the predictors of out-migration. If available, we will also look at observed characteristics from the later survey rounds to compare those who stayed in the sample vs those who in-migrated into the study area.

In the original paper, the authors allow those who migrated in and out of the study to continue to contribute person-time even if they had out-migrated. We will right censor people at their first instance of out-migration. After right censoring, we will generate new incidence rate estimates and examine how migration may have affected the null result. We will also use survival analysis to assess the hazard of developing HIV incidence after right-censoring.

Using a competing risks model, we will assess if the association between treatment group and HIV incidence changed after accounting for migration.

### *3.2.2.4 Subsample analysis by location*

The authors' note that location matters in terms of HIV prevalence is an interesting development that deserves some study. Being nearer to a highway would theoretically increase the probability of more sexual partners, increasing the likelihood of viral transmission. As a result, these individuals are likely compositionally different from others in the study.

We will look first to determine whether there are compositional differences between these areas. Then, we will examine incidence rates between areas with a highway and areas without. If available, the distance of a community from a major highway could be an important predictor of HIV incidence, meaning that programs should target areas connected to population centres. An indicator for the distance of a community from a major highway will be included in the model to assess its association with HIV incidence.

## **4. Conclusion**

In this study, we propose to replicate Iwuji et al. (2018) (ANRS 12249). The study was the first of four cluster randomized trials aimed at understanding whether treatment as prevention (TasP) programs were effective at the population level, rather than at the sample level as seen in Cohen et al. (2011) (HPTN 052). The original authors show that TasP in rural South Africa was ineffective due to poor linkage of care and high in- and out-migration

in the study area. Our plan first conducts a pure replication, using the authors' data and methods to ensure that they are valid. We will then test the result further, to see if we can highlight further reasons beyond the null result. The first of these will be to use survival modelling to assess the impact of the treatment on HIV transmission time. Second, we will test how the CD4 cut-off change affected the control group and overall incidence rates. Third, we will then account for the high in- and out-migration in the study to see how increased migration may have contributed to the null result. Finally, we will look at location to see if there is a difference in HIV incidence based on proximity to major highways. Answers to these questions will deepen the understanding of the null result, and provide policy- and decision-makers a clearer picture of how to make TasP programs more effective in the future.

## References

- Austin, P. C. (2017). A tutorial on multilevel survival analysis: Methods, models and applications. *International Statistical Review*, 85(2), 185–203. <https://doi.org/10.1111/insr.12214>
- Bärnighausen, T., Eyal, N., & Wikler, D. (2014). HIV Treatment-as-Prevention Research at a Crossroads. *PLoS Medicine*, 11(6), 1–4. <https://doi.org/10.1371/journal.pmed.1001654>
- Brown, A. N., Cameron, D. B., & Wood, B. D. K. (2014). Quality evidence for policymaking: I'll believe it when I see the replication. *Journal of Development Effectiveness*, 6(3), 215–235. <https://doi.org/10.1080/19439342.2014.944555>
- Cohen, M. S., Chen, Y. Q., Mccauley, M., Gamble, T., Hosseinipour, M. C., Kumarasamy, N., ... Fleming, T. R. (2011). Prevention of HIV-1 infection with early antiretroviral therapy. *New England Journal of Medicine*, 365(6), 493–505. <https://doi.org/10.1056/NEJMoa1105243>
- Department of Health, R. of S. A. (2014). *National consolidated guidelines for the prevention of mother-to-child transmission of HIV (PMTCT) and the management of HIV in children, adolescents and adults*.
- Havlir, D. V., Balzer, L. B., Charlebois, E. D., Clark, T. D., Kwarisiima, D., Ayieko, J., ... Petersen, M. (2019). HIV Testing and Treatment with the Use of a Community Health Approach in Rural Africa. *New England Journal of Medicine*, 381(3), 219–229. <https://doi.org/10.1056/NEJMoa1809866>
- Hayes, R., Ayles, H., Beyers, N., Sabapathy, K., Floyd, S., Shanaube, K., ... White, R. (2014). HPTN 071 (PopART): Rationale and design of a cluster-randomised trial of the population impact of an HIV combination prevention intervention including universal testing and treatment - a study protocol for a cluster randomised trial. *Trials*, 15(1), 1–17. <https://doi.org/10.1186/1745-6215-15-57>
- Iwuji, C. C., Orne-Gliemann, J., Larmarange, J., Balestre, E., Thiebaut, R., Tanser, F., ... Hontelez, J. (2018). Universal test and treat and the HIV epidemic in rural South Africa: a phase 4, open-label, community cluster randomised trial. *The Lancet HIV*, 5(3), e116–e125. [https://doi.org/10.1016/S2352-3018\(17\)30205-9](https://doi.org/10.1016/S2352-3018(17)30205-9)
- Larmarange, J., Diallo, M.H., McGrath, N., Iwuji, C., Plazy, M., Thiebaut, R., ... Orne-Gliemann, J. (2018). The impact of population dynamics on the population HIV care cascade: results from the ANRS 12249 Treatment as Prevention trial in rural KwaZulu-Natal (South Africa). *Journal of the International AIDS Society*, 21(S4), e25128. <https://doi.org/10.1002/jia2.25128>
- Moore, D. J., Rooney, A., Gouaux, B., Georges, S., Depp, C. A., Atkinson, J. H., ... Blackstone, K. (2013). Preliminary evidence for feasibility, use, and acceptability of individualized texting for adherence building for antiretroviral adherence and substance use assessment among HIV-infected methamphetamine users. *AIDS Research and Treatment*, 2013. <https://doi.org/10.1155/2013/585143>
- Rodger, A. J., Cambiano, V., Bruun, T., Vernazza, P., Collins, S., Van Lunzen, J., ... Lundgren, J. (2016). Sexual activity without condoms and risk of HIV transmission in serodifferent couples when the HIV-positive partner is using suppressive antiretroviral therapy. *JAMA - Journal of the American Medical Association*, 316(2), 171–181. <https://doi.org/10.1001/jama.2016.5148>
- Tanser, F., Bärnighausen, T., Cooke, G. S., & Newell, M. L. (2009). Localized spatial clustering of HIV infections in a widely disseminated rural South African epidemic. *International Journal of Epidemiology*, 38(4), 1008–1016. <https://doi.org/10.1093/ije/dyp148>