AIDS and Behavior (2024) 28:3873–3882
https://doi.org/10.1007/s10461-024-04463-3

ORIGINAL PAPER

Efficacy of the Dapivirine Vaginal Ring Accounting for Imperfect Adherence

Marla J. Husnik[1,2] · Renee Heffron[1,3] · James P. Hughes[4] · Barbra Richardson[4] · Ariane van der Straten[5,14] · Thesla Palanee-Phillips[6] · Lydia Soto-Torres[7] · Devika Singh[8,10] · Brenda Gati Mirembe[9] · Edward Livant[10] · Zakir Gaffoor[11] · Leila E. Mansoor[12] · Samantha S. Siva[11] · Sufia Dadabhai[13] · Flavia Matovu Kiweewa[9] · Jared M. Baeten[1,3] · for the MTN-020/ASPIRE Study Team

Accepted: 2 August 2024 / Published online: 19 August 2024

Abstract
Product adherence is critical to obtaining objective estimates of efficacy of pre-exposure prophylactic interventions against HIV-1 infection. With imperfect adherence, intention-to-treat analyses assess the collective effects of complete, sub-optimal and non-adherence, providing a biased and attenuated estimate of the average causal effect of an intervention. Using data from the MTN-020/ASPIRE phase III trial evaluating HIV-1 efficacy of the dapivirine vaginal ring, we conducted per-protocol, and adherence-adjusted causal inference analyses using principal stratification and marginal structural models. We constructed two adherence cut offs of ≥ 0.9 mg (low cutoff) and > 4.0 mg (high cutoff) that represent drug released from the ring over a 28-day period. The HIV-1 efficacy estimate (95% CI) was 30.8% (3.6%, 50.3%) (P = 0.03) from the per-protocol analysis, and 53.6% (16.5%, 74.3%) (P = 0.01) among the highest predicted adherers from principal stratification analyses using the low cutoff. Marginal structural models produced efficacy estimates (95% CIs) ranging from 48.8 (21.8, 66.4) (P = 0.0019) to 56.5% (32.8%, 71.9%) (P = 0.0002). Application of adherence-adjusted causal inference methods are useful in interpreting HIV-1 efficacy in secondary analyses of PrEP clinical trials.

Keywords Efficacy · Non-adherence · HIV-1 prevention · Dapivirine · Intravaginal ring · Causal inference

The members of the "for the MTN-020/ASPIRE Study Team" are listed in acknowledgements section.

✉ Jared M. Baeten
jbaeten@uw.edu

1 Department of Epidemiology, University of Washington, Seattle, WA, USA

2 Statistical Center for HIV/AIDS Research and Prevention, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

3 Department of Global Health, University of Washington, Seattle, WA, USA

4 Department of Biostatistics, University of Washington, Seattle, WA, USA

5 Department of Medicine, The Center for AIDS Prevention Studies, University of California San Francisco, San Francisco, CA, USA

6 Faculty of Health Sciences, Wits Reproductive Health and HIV Institute, University of the Witwatersrand, Johannesburg, South Africa

7 National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA

8 University of Pittsburgh, Pittsburgh, PA, USA

9 Makerere University John Hopkins University Research Collaboration (MU-JHU), Kampala, Uganda

10 Magee-Womens Research Institute, Pittsburgh, PA, USA

11 HIV Prevention Research Unit, South African Medical Research Council, Cape Town, South Africa

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

13 Johns Hopkins Bloomberg School of Public Health, Blantyre, Malawi

14 ASTRA Consulting, Kensington, CA, USA

Background

Obtaining unbiased estimates of efficacy of pre-exposure prophylactic (PrEP) interventions against HIV-1 infection relies on objective measures of product adherence. In clinical trials, an increasing number of PrEP strategies have been shown to be effective at reducing risk of HIV-1 infection, including evidence of very high HIV-1 protection when used with high adherence [1–3]. However, several clinical trials [4–6] that studied the use of oral tablets and topical vaginally delivered PrEP among women in sub-Saharan Africa failed to show high efficacy due to sub-optimal adherence. In the MTN-020/ASPIRE trial of the dapivirine vaginal ring, intention-to-treat (ITT) analyses showed a 27% reduction in HIV-1 risk overall; this increased to 37% after removing data from participants in two trial sites contributing to lower adherence and to 56% among women > 21 years of age, who showed evidence of better adherence compared to their younger counterparts [7].

ITT analyses are typically implemented to estimate the average causal effect (ACE) in interventional studies. However, in the presence of imperfect visit retention and/or product adherence, ITT analyses are limited to assessing a combined effect of adherers, partial adherers, and non-adherers, and thus provide a biased and attenuated estimate of the ACE of an intervention [8]. Many trials incorporate per-protocol analyses into their analysis plans, which estimate the ACE only over the time when PrEP was actually received; however, such analyses could also be biased due to informative censoring or selection bias due to suboptimal product use.

Additional analytical approaches have aimed to provide unbiased estimates of efficacy in the context of incomplete follow-up and/or non-adherence to product use. These include the causal inference methods of principal stratification and marginal structural models utilizing inverse-probability-of-censoring weights (IPCW) and/or inverse-probability-of-treatment weights (IPTW). Principal stratification estimates efficacy among varying levels of adherers and allows for estimates of the complier average causal effect (CACE). Marginal structural models estimate efficacy among high adherers adjusting for time-varying confounding and selection bias. We applied these methods to the MTN-020/ASPIRE study data to improve understanding of HIV-1 efficacy from the dapivirine vaginal ring related to high adherence and explore limitations of these methods to fully account for non-adherence.

Methods

MTN-020/ASPIRE was a fifteen-site, randomized, double-blind, placebo-controlled, phase III trial designed to test the safety and effectiveness of monthly use of an intravaginal ring releasing dapivirine, a non-nucleoside reverse transcriptase inhibitor (NNRTI). The trial was carried out in Malawi, South Africa, Uganda, and Zimbabwe from August 2012 through June 2015. Details of the design, methods, and results of the trial (ClinicalTrails.gov number NCT01617096) are reported elsewhere [7, 9]. Institutional review boards at each participating clinical site approved the study protocol, and participants provided written informed consent for participation and specimen storage.

Population and Study Procedures

Sexually active, non-pregnant, HIV-1 seronegative women between the ages of 18–45 years were randomized (1:1) to use a ring containing 25 mg of dapivirine or a matching placebo ring. At monthly visits, women were instructed to insert and wear the ring continuously until their next visit when a new ring would be inserted. Clinic staff collected baseline demographic, and monthly sexual risk behavior, male partner characteristics, menstrual bleeding, vaginal practices, and adherence data on case reports forms. Testing for sexually transmitted infections including chlamydia, gonorrhea, trichomoniasis, and syphilis was done at baseline and post-randomization. Blood specimens were collected for HIV-1 serologic testing monthly and incident HIV-1 infections were adjudicated by an endpoints committee blinded to treatment assignment and to adherence to product use of the study participants.

Covariates Used in Models

Baseline covariates assessed as predictors of adherence in logistic regression models for principal stratification, and inverse-probability of censoring/treatment weighted marginal structural models were as follows: study site, age (18–21, 22–26, 27–45 years old), marital status (not married, married), highest level of education (no secondary education, some secondary education or higher), alcoholic drinks per week (none, 1–6, 7+), owns mobile phone (yes/no), travel time to clinic (< 30 min, 30–60 min, 1–2 h, > 2 h), earns own income (yes/no), body mass index (underweight: ≤ 18.5, normal weight: 18.5–24.9, overweight: 25–29.9, obesity: ≥ 30), number of vaginal sex acts in past 3 months, transactional sex in the past year (no, yes), anal sex (no, yes), any bacterial sexually transmitted infections, worried about having a vaginal ring inside of her every day for at least a year (not at all, somewhat, very worried). Covariates during follow-up included the following: percentage of vaginal sex acts protected by a condom in past 7 days [0%, 1–49%, 50–99%, 100% (or no sex)], male or female condom use during last act of vaginal sex in past 7 days, anal sex in past 3 months (no, yes), number of male sex partners in past 3 months (0–1, 2, 3 or more), sex with a primary partner in past 3 months (no, yes, no primary partner), primary partner knows participant is taking part in the study (no/not sure, yes, no primary partner), primary partner knows participant is using vaginal ring (no/unsure, yes, no primary partner), same primary partner in past 3 months (no, yes, no primary partner), HIV status of primary partner in past 3 months (negative, positive, don't know, no primary partner), primary partner is taking antiretroviral drugs (no, yes, don't know, no primary partner), primary partner is circumcised (no, yes, don't know, no primary partner), bothered wearing ring every day (no, yes), number of times started or had her menstrual period in the last 3 months (0, 1, 2, 3), used something to control spotting or bleeding in last 3 months (no, no spotting/bleeding, yes), put anything inside vagina in last 3 months (no, yes), primary family planning methods [implants, intrauterine device, injectables: norethisterone enanthate and intramuscular depot medroxyprogesterone acetate, other (includes oral contraceptive pills, sterilization, male or female condoms only, no family planning method)].

Measures of Adherence

Returned vaginal rings were shipped to a central laboratory and tested for residual dapivirine using acetone extraction and high-pressure liquid chromatography [10]. Vaginal rings from participants randomized to the active group only were used to construct two adherence cut offs of ≥ 0.9 mg and > 4.0 mg representing the amount of drug released from the ring over the expected 28-day monthly visit period. Manufacturing load levels minus residual drug levels were used to calculate the amount of drug released from each ring. Since study visits may not have been spaced exactly to 28 days, the ratio of the amount of drug released to the number of days since dispensation was calculated and normalized to estimate the amount of dapivirine released during a standard 28-day period. A cut-off of < 0.9 mg was chosen to represent no or very low use of the ring, equating to one standard deviation of lab measurement error > 0 mg dapivirine released based on testing of unused rings. We consequently defined an adherence outcome of any use of the ring at a level of ≥ 0.9 mg drug released during a 28-day period. In two phase I trials and one phase I/II trial [11–13], investigators concluded that an average of 4.0–5.0 mg of dapivirine was released over a 28-day period of continuous ring use, and thus in this study > 4.0 mg was also chosen as a more stringent measure, approximating continuous use [14].

Application of Statistical Methods

We apply three statistical methods (per-protocol, principal stratification, and marginal structural models) to a robust clinical trial dataset from the MTN-020/ASPIRE study utilizing objective measures of adherence to the PrEP agent (dapivirine) within a vaginal ring. The per-protocol analysis provides the net treatment effect of the active drug adjusted for treatment assignment. The principal stratification analysis provides an estimate of the CACE. Empirical estimates of HIV efficacy were calculated for comparison purposes. The marginal structural model analysis utilizes two IPCWs and an IPTW to provide an estimate of the ACE if all participants were to be continuously compliant to product use throughout the study.

Per-protocol Analysis

In per-protocol analyses, we constructed two Cox proportional hazards (PH) models, each stratified by site with treatment arm (either dapivirine or placebo ring) included as the predictor. One of the models included additional covariates previously found to be statistically significant predictors of HIV-1 infection in subgroup analyses and so were used to control for confounding [7]. Time to HIV-1 seroconversion was defined to mimic participant follow-up time as if accrued under perfect protocol adherence, censoring at the last negative HIV test before detection of pregnancy (whereby participants were put on product hold) or the first occurrence of 3+ non-compliant events in a 12 month period that included the following: a missed visit, participants reporting the ring being out for more than 12 h, and participant failing to return product to the clinical site. The event time was based on the time from enrollment to the earliest time of the following: last negative HIV test prior to when the participant was determined to be pregnant, last negative HIV test prior to the first occurrence of three or more non-compliant events in a 12-month period as stated above, HIV infection, or the end of study follow-up. The censor indicator was then based on the participant HIV-1 status at this newly defined event time. This analysis tested the hypothesis that there is no treatment effect using maximum likelihood estimates and the score test. Estimates of the percent efficacy and 95% confidence intervals (CIs) are reported based on 1-hazard ratio obtained from Cox PH models.

Principal Stratification Analysis

We use the potential outcomes framework to define "principal strata" that are independent of the treatment assignment. The principal strata adjust for characteristics inherent in the post-randomization variable of adherence without introducing selection bias as might be the case under a per-protocol analysis [15]. This method assumes the exclusion restriction for identifiability [16]. To differentiate adherers from non-adherers, we used the two measures of ring adherence of ≥ 0.9 mg and > 4.0 mg of drug released over 28 days. Using the 6-month dataset, a time-point providing the most participant-level data over study follow-up, and 12-month for sensitivity analyses, we constructed a logistic regression model using participant baseline characteristics among active arm participants only since adherence as defined by the amount of dapivirine released from the ring can only be measured in the active arm. We followed Shtatland, Kleinman, and Cain for finding the best predictive model utilizing a three-step process [17]. We used the logistic regression models to obtain the predicted probabilities of adherence among all participants in the ASPIRE cohort. Finally, we constructed a Cox PH model with treatment arm, the predicted probability of adherence and an interaction term between treatment arm and the predicted adherence level [18]. We calculated hazard ratios and 95% CIs to obtain percent efficacy estimates at the median values of three strata defined by predicted probabilities of adherence of 0.86–0.90, 0.91–0.95, and 0.96–1.00, levels considered in this context to be relevant as treatment compliers.

Marginal Structural Models

The goal for this analysis was to eliminate informative censoring and selection bias, that occur when participants are non-adherent for reasons associated with being in the study, by creating marginal structural models [19–22]. We constructed two IPCWs based on censoring due to pregnancy, and due to informative loss to follow-up (i.e., the participant was terminated for one of the following reasons: investigator decision, refused further participation, unable to contact participant, or relocated with no follow-up planned). To account for time-varying confounding of exposure to treatment we constructed an additional weight based on the adherence outcome of ≥ 0.9 mg of released drug in the active arm. Because ring adherence data did not exist in the placebo arm, the value of 1.0 was assigned as the IPTW for placebo participants. This corresponds to the average IPTW among the active arm participants. Assumptions for marginal structural models are as follows: (1) consistency, (2) exchangeability, (3) positivity, and (4) no misspecifications of the models [19, 22]. We constructed non-truncated and truncated (removing weights below the 1%tile, and above the 99th%tile) stabilized IPCWs and IPTW [23] that were theoretically proportional to a subject receiving her own exposure history using pooled logistic regression. A composite weight was created by multiplying the IPCWs and the IPTW. Final weighted pooled logistic regression models were then constructed using only the IPTW and the composite weight to estimate the effect of dapivirine use. All analyses excluded participant-visits in the first year of the trial owing to missing residual drug levels in vaginal rings preventing estimation of weight models. All models censor participants at the time of first pregnancy (N = 176) and exclude 37 primary endpoints due to missing data in IPCW and IPTW models. Behavioral data collected on a quarterly visit schedule was aligned with the monthly ring-use data utilizing the last value carried forward convention where appropriate. Analyses were conducted in SAS, version 9.4 (SAS Institute, Cary, North Carolina).

Results

There were 2629 participants randomized into the study, however 15 participants dropped out prior to the first visit where HIV-1 testing occurred (Table 1). Therefore, a total of 2614, 1308 in the dapivirine arm and 1306 in the placebo arm, were included in analyses. Median age was 27 years, over 80% had completed secondary school or higher, ~ 40% were married, the majority did not drink alcohol, over 90% owned a cell phone, and roughly half earned their own income.

Table 2 shows the results from the per-protocol analyses. Model 1 was stratified by site and included only the randomization arm as a predictor for HIV seroconversion. The estimated percent efficacy was 28.8% (95% CI 0.9%, 48.9%; $P = 0.04$), and with additional adjustment for confounding was 30.8% (95% CI 3.6%, 50.3%; $P = 0.03$). These estimates were very close to those obtained from the ITT analyses of 27% (95% CI 1%, 46%; $P = 0.05$) [7].

Table 3 and Fig. 1 show the results of the principal stratification analyses. HIV-1 efficacy results for the first set of models utilizing month-6 data to obtain the predicted probabilities of adherence of 0.86–0.90, 0.91–0.95, and 0.96–1.00 using the ≥ 0.9 mg cut off to define adherence were all statistically significant at the $P < 0.05$ level and ranged from 29.7 (95% CI 3.3%, 48.4%) to 53.9% (95% CI 16.5%, 74.8%). Empirical estimates ranged from 22.3 (95% CI 15.5%, 30.4%) to 48.2 (95% CI 42.2%, 54.8%). For the 12-month dataset using the same cut off for compliance, HIV-1 efficacy estimates ranged from 30.8 (95% CI 3.3%, 50.6%) to 40.7% (− 17.6%, 70.4%). Empirical estimates ranged from 14.2 (95% CI 12.0%, 17.0%) to 59.9% (95% CI 43.4%, 82.2%). Using 6-month data and the > 4.0 mg cut off for adherence we found one, albeit unstable estimate, for HIV-1 efficacy in the 0.86–0.90 predicted adherence group of 39.6% (95% CI − 3517.8%, 99.0%). Due to the sparsity of information for the remaining adherence groups using the 6-month dataset, and also when using the 12-month dataset, the estimated predicted probabilities and hazard ratios were inestimable in those regions.

For analyses utilizing IPCW and IPTW to construct marginal structural models, results are shown in Table 4. Models 1 and 3 show estimated percent HIV-1 efficacy using non-truncated stabilized weights for adherence alone, and the composite of adherence, first pregnancy and informed loss to follow-up, respectively. Models 2 and 4 show estimates using stabilized weights truncated at 1%tile and 99th%tiles for adherence alone, and the composite weight, respectively.

Table 1 Baseline characteristics of MTN-020/ASPIRE participants

Characteristic	Dapivirine ring N = 1308 N (%)	Placebo ring N = 1306 N (%)
Age		
18–21	269 (20.6)	242 (18.5)
22–26	396 (30.3)	445 (34.1)
≥ 27	643 (49.2)	619 (47.4)
Education		
< Secondary school	212 (16.2)	191 (14.6)
≥ Secondary school or higher	1096 (83.8)	1115 (85.4)
Married	525 (40.1)	544 (41.7)
Alcohol drinks per week		
None	1137 (87.0)	1158 (88.7)
1–6	139 (10.6)	120 (9.2)
7 or more	31 (2.4)	28 (2.1)
Owns a mobile phone	1179 (90.1)	1187 (90.9)
Earns own income	604 (46.2)	577 (44.2)
Transactional sex[a]	73 (5.6)	88 (6.8)
Worried about having ring inside her every day for at least a year?		
Not at all worried	921 (70.4)	918 (70.3)
Somewhat worried	366 (28.0)	370 (28.3)
Very worried	21 (1.6)	18 (1.4)
Sexually transmitted infections[b]	284 (21.7)	259 (19.8)
Condom-protected vaginal sex in the last 7 days (no sex = 100% protected)		
0%	348 (26.6)	398 (30.5)
1–49%	64 (4.9)	67 (5.1)
50–99%	104 (8.0)	84 (6.4)
100%	792 (60.6)	756 (57.9)
Number of sex partners		
0–1	1097 (83.9)	1082 (82.9)
2 partners	156 (11.9)	147 (11.3)
3+ partners	55 (4.2)	76 (5.8)
Primary partner knows ppt. is taking part in study		
No/not sure	330 (25.2)	309 (23.7)
Yes	972 (74.3)	990 (75.9)
No primary partner	6 (0.5)	6 (0.5)

[a]The question asked on the case report form was: "In the past year, did you receive money, material goods, gifts, drugs, or shelter in exchange for vaginal or anal sex?"

[b]Positive for any one of the following: chlamydia, gonorrhea, trichomoniasis, or syphilis

HIV-1 efficacy (95% CIs) across these four marginal structural models ranged from 48.8 (21.8%, 66.4%) (Model 2) to 56.5% (32.8%, 71.9%) (Model 3). All estimates were statistically significant at levels of $P < 0.002$.

Discussion

In this analysis of data from the MTN-020/ASPIRE trial of the dapivirine vaginal ring, we used multiple analytical

Table 2 Results of Cox proportional hazards models for per-protocol analyses

Cox proportional hazards model	Estimated percent HIV-1 efficacy	Estimated percent HIV-1 efficacy 95% CI	P value
Model 1[a]	28.8	0.9, 48.9	0.04
Model 2[b]	30.8	3.6, 50.3	0.03

[a]This model was stratified by site and included randomization arm

[b]This model was stratified by site and included randomization arm, and additionally the following variables to adjust for potential confounding: [age group (<25, ≥25 years old), marital status (married, not married), number of male sex partners (≥2 sex partners in last 3 months, 0–1 sex partners in last 3 months), sexually transmitted infections (any positive baseline result, all negative baseline results for the following: chlamydia, gonorrhea, trichomoniasis, or syphilis)

methods that account for imperfect protocol and product adherence to estimate HIV-1 protection. Per-protocol analyses, which aimed to mimic perfect protocol adherence, although without explicit consideration of product use, yielded efficacy estimates of approximately 30%, very near the ITT finding of 27% for this trial [7]. Causal models, specifically principal stratification and marginal structural, which considered both protocol and product adherence, estimated HIV-1 protection to be between 41% and 57%.

All analysis approaches we evaluated generated HIV-1 efficacy estimates that were higher compared to ITT, which was expected, given that retention and product adherence were known to be imperfect in the trial. Per-protocol analyses corrected for adherence by censoring at the time of first missed visits producing efficacy estimates only a few percentage points higher than those found in the ITT analysis. Such analyses did not account for selection bias due to informative censoring and assumed participants would have been fully compliant with product use. Principal stratification analyses aimed to categorize participants into sub-groups defined by predicted probabilities of adherence based on enrollment characteristics, although assignment into strata that clearly defined highly adherent groups was difficult in this dataset due to the lack of strong predictors of

Table 3 Principal Stratification: Estimated and empirical HIV-1 efficacy for predicted probability of adherence groups

Visit month[a]	Adherence cut-off (≥0.9 mg & >4.0 mg) (mg)	Number of participants, person-months of follow-up, number of HIV-1 endpoints	Predicted probability of adherence[b]	Empirical HIV-1 percent efficacy[c]	Empirical HIV-1 percent efficacy 95% CI	Estimated HIV-1 percent efficacy[d]	Estimated HIV-1 percent efficacy 95% CI	P value
6	≥0.9	446, 9277, 34	0.86–0.90	22.3	15.5, 30.4	29.7	3.3, 48.4	0.03
6	≥0.9	1375, 28,615, 63	0.91–0.95	48.2	42.2, 54.8	42.9	13.2, 61.6	0.008
6	≥0.9	143, 2858, 12	0.96–1.00	30.0	20.7, 43.5	53.9	16.5, 74.8	0.01
12	≥0.9	557, 12,262, 47	0.86–0.90	14.2	12.0, 17.0	30.8	3.3, 50.6	0.03
12	≥0.9	692, 14,526, 32	0.91–0.95	48.1	36.9, 60.3	36.3	− 4.4, 60.5	0.07
12	≥0.9	279, 5840, 6	0.96–1.00	59.9	43.4, 82.2	40.7	− 17.6, 70.4	0.14
6	>4.0	7, 144, 0	0.86–0.90	Inestimable[e]	Inestimable[e]	Unstable estimate[f]	Unstable estimate[f]	–
6	>4.0	–	0.91–0.95	–	–	Inestimable[g]	Inestimable[g]	–
6	>4.0	–	0.96–1.00	–	–	Inestimable[g]	Inestimable[g]	–
12	>4.0	–	0.86–0.90	–	–	Inestimable[g]	Inestimable[g]	–
12	>4.0	–	0.91–0.95	–	–	Inestimable[g]	Inestimable[g]	–
12	>4.0	–	0.96–1.00	–	–	Inestimable[g]	Inestimable[g]	–

[a]Data from this visit month was used to construct the logistic regression models needed to obtain predicted probabilities of adherence based on best model using the Akaike Information Criterion (AIC). The number of participants included was as follows: 6-month model N = 702, 12-month model N = 1113

[b]Predicted probabilities were obtained through logistic regression models using the adherence cut-off as outcome and baseline covariates as predictors

[c]Garwood F (1936). Fiducial Limits for the Poisson Distribution. Biometrika 46:441–453

[d]Efficacy estimates based on Cox proportional hazards models including arm, adherence group and arm x adherence group interaction term

[e]There were 0 endpoints in each arm, so therefore the empirical estimate was inestimable

[f]The HIV-1 percent efficacy for this adherence group based on the Cox proportional hazards model was: 39.6 (95% CI − 3517.8, 99.0; P = 0.81)

[g]No participants had predicted probabilities of adherence at these levels and thus HIV-1 percent efficacy (95% CIs) was inestimable

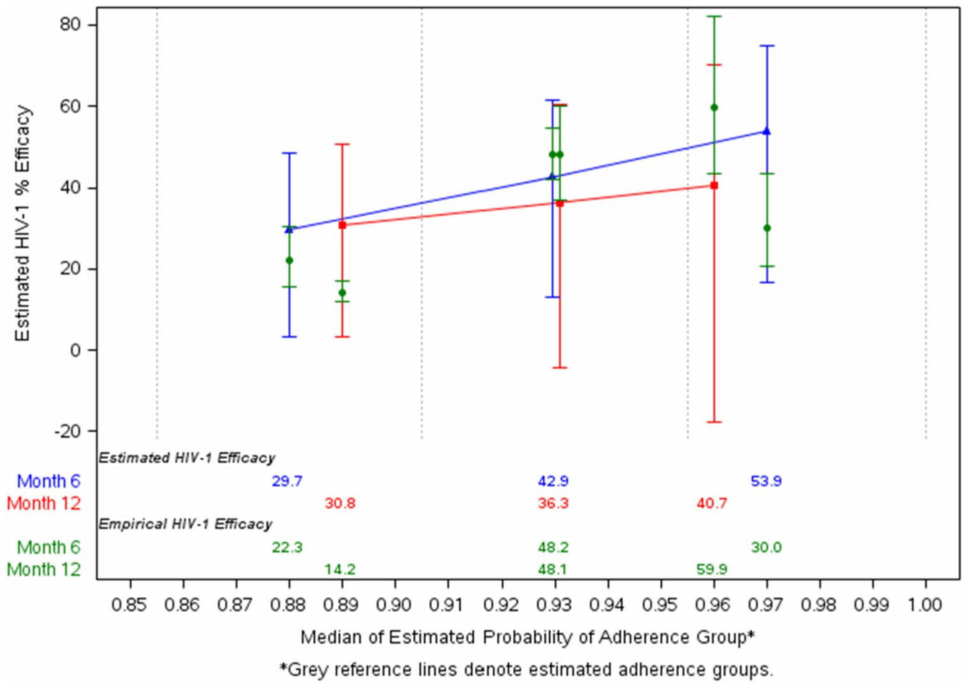

Fig. 1 Principal Stratification: Estimated and empirical HIV-1 efficacy for predicted probability of adherence groups

Table 4 Results of marginal structural models

Generalized estimating equation (GEE) model	Number of HIV-1 endpoints	Estimated percent HIV-1 efficacy	Estimated percent HIV-1 efficacy 95% CI	P value
Model 1[a]	131	49.0	22.1, 66.6	0.0018
Model 2[b]	131	48.8	21.8, 66.4	0.0019
Model 3[c]	131	56.5	32.8, 71.9	0.0002
Model 4[d]	131	54.3	30.0, 70.1	0.0003

[a]Using non-truncated stabilized weights for predicting adherence (in active arm only & using ≥ 0.9 mg cut-off), censored at time of first pregnancy

[b]Using truncated stabilized weights (at 1%tile and 99th%tile) for predicting adherence (in active arm only & using ≥ 0.9 mg cutoff), censored at time of first pregnancy

[c]Using non-truncated stabilized weights for predicting adherence (in active arm only & using ≥ 0.9 mg cut-off), time to first pregnancy and informative loss to follow-up. Data set was censored at time of first pregnancy

[d]Using truncated stabilized weights (at 1%tile and 99th%tile) for predicting adherence (in active arm only & using ≥ 0.9 mg cutoff), time to first pregnancy and informative loss to follow-up. Data set was censored at time of first pregnancy

adherence. The two residual drug release cut points to define a participant as adherent were either arguably not stringent enough (≥ 0.9 mg) or overly stringent as evidenced by scarcity of data points (> 4.0 mg). A study with a point-exposure instead of one that is time-dependent and user-dependent (as far as it affects consistency or misclassification of exposure), would have more likely met assumptions needed for marginal structural modelling. Compared to the other models, however, the marginal structural models likely provided a closer estimate of true efficacy given it reweighted the dataset into a pseudo-population to account for time-varying confounding and selection bias at the same time.

By creating strata of exchangeable participants [24, 25] considered to be moderately-high adherers (86–90%), and strata considered to be very-high adherers (91–95% and 96–100%) when using the ≥ 0.9 mg cut off for adherence, we found suggestion of monotonic increasing HIV-1 efficacy estimates for each of the 6- and 12-month datasets. Notably, the empirical estimates and their 95% CIs trended similarly and were also contained within the 95% CIs of the predicted estimates. This is likely due to the additional statistical uncertainty from inclusion of the adherence parameter in the estimated models. Likewise, estimates of ring efficacy in the lower adherent group (86–90%) had comparable estimates to those seen in ITT and per-protocol analyses. However, in the most adherent strata (96–100%), we observed HIV-1 efficacy estimates ranging from 41 to 54%, albeit scarcity in the endpoints for the more stringent adherence groups resulted in highly imprecise or inestimable estimates. In marginal structural models, however, precision of efficacy estimates was not an issue, due to the large number of endpoints included in each model [although 22% (37/168) of endpoints were excluded from the analysis due to the unavailability of residual drug data from rings in the first year of the MTN-020/ASPIRE study]. The estimates ranged from 49% to 57%, depending on whether stabilized IPTC/IPCW's were truncated, and whether weights were used for adherence alone, or a combined weight adjusting for adherence, first pregnancy or informative loss-to-follow-up.

Principal stratification analyses included two definitions of adherence based on release of drug in vaginal rings. Considering the principal stratification efficacy estimates and their 95% confidence intervals using 6-month data, the ≥ 0.9 mg drug release definition of adherence yielded estimates whereas those of the > 4.0 mg drug release definition were inestimable due to data scarcity. This is not surprising since a less stringent definition of adherence includes a greater proportion of participants categorized as adherent, although likely misclassifies participants as adherent when they are not. The effect of exposure misclassification is typically attenuation of efficacy estimates. However, the group whose predicted probability of adherence is 96–100% likely contains participants who have the highest probability of adhering for the lower adherence threshold. Therefore, an efficacy estimate of 53.6% (95% CI 16.5%, 74.3%) may provide a reasonable estimate of the CACE, since we can be more confident in this strata that the exclusion restriction for identifiability assumption is being met. This assumption is based on the notion that potential non-adherers do not derive a treatment effect [15]; in our case, participants whose predicted probability of being adherent is very high at a lower threshold for adherence, is unlikely to include consistently non-adherent individuals given a cut point of < 0.9 mg represents no or very low use of the ring. Likewise, this estimate is like that seen in subgroup analyses of women > 21 years of age with an efficacy estimate of 56%, and who had evidence of better adherence compared to their younger counterparts. Estimates at 12 months serving as sensitivity analyses, showed attenuated point estimates although confidence intervals overlapped with the 6-month measures. Of note, in the highest predicted probability of adherence group (0.96–1.00) there were relatively fewer HIV-1 endpoints (1.0 per 1000 person-months at 12 months vs. 4.2 per 1000 person-months at 6 months) possibly due in part to a study-wide temporal effect that was seen where adherence increased and began to level off at 12 months into the trial [26], and thus numbers of HIV-1 seroconversions began to decrease due to increased adherence in both arms. Also, removing the first year of ring data from this analysis may have influenced the prediction models for categorizing participants into predicted adherence groups.

If assumptions of the marginal structural models are met, then the exposure is independent from the measured confounders in the pseudo-population. The consistency assumption asserts that a participant's exposure history is well-defined such that the potential outcome of a participant's exposure history is exactly that of the observed outcome [27]. However, in the case of exposure to the dapivirine ring, this assumption is clearly violated. As shown in the principal stratification analyses, we observed that inherent to the estimation of the CACE is determination of an appropriate cut point for adherence. Misclassification bias is present since defining a perfect cut point, based on drug release to define sustained exposure to the ring that warrants protection, is unknown. Therefore, we are left with using cut-points of the amount of drug released that are considered reasonable to define any and continuous use. However, a preferable definition to meet the consistency assumption would be one where exposure to the ring was completely controlled by the investigator such that consistent and sustained exposure over a 28-day period is guaranteed. Without this, one cannot confidently assume that the consistency assumption has been met. Second, the assumption of exchangeability, or no unmeasured confounding, may have also been violated. Although we included all potential confounders collected during the study when constructing the IPCW and IPTW models, we cannot be completely confident other participant or study characteristics did not also play a role in adherence, pregnancy or loss-to-follow-up, and the outcome of HIV-1 infection. For example, other studies based on the MTN-020/ASPIRE cohort have shown that adherence improved over time when participants adjusted to wearing the ring and their fears of negative side effects dissipated [28], or that participants initially were concerned about the novelty of the ring but over time developed a sense of comfort and acceptability increased [13, 29–31]. Although a variable pertaining to worries about ring use was included in the models, we did not adjust specifically for changes in perception or fears of use over time, which, if also related to the outcome of HIV-1 infection, could be considered an unmeasured confounder, or at least one with residual confounding. The additional assumption of positivity is reasonable since the MTN-020/ASPIRE cohort was large enough that over the distribution of covariates in the weight models, there were participants that did and did not have the outcomes (adherence, pregnancy, and informative loss to follow-up). Finally, as in any modeling exercise, it could be the case that there was misspecification of models for baseline and time-varying covariates on adherence to ring use, pregnancy, and informative loss to follow-up as well as the final marginal structural models for effect of adherence to ring use on HIV-1 infection, controlling for baseline covariates.

Despite several limitations, the application of statistical methods utilizing the potential outcomes framework that accounted for non-compliance produced higher HIV-1 efficacy estimates than the ITT analysis. Obtaining estimation of efficacy of PrEP strategies such as the use of a vaginal ring for and among high-adherers is valuable for planning and implementation of future HIV-1 prevention studies, and rollout.

Acknowledgements We wish to thank the MTN-020/ASPIRE study team, MTN leadership and MTN Laboratory Center staff, as well as the dedication of all study participants for making this research possible.

MTN-020/ASPIRE STUDY TEAM Study Team Leadership: Jared Baeten, University of Washington (Protocol Chair); Thesla Palanee-Phillips, Wits Reproductive Health and HIV Institute (Protocol Co-chair); Elizabeth Brown, Fred Hutchinson Cancer Research Center (Protocol Statistician); Lydia Soto-Torres, US National Institute of Allergy and Infectious Diseases (Medical Officer); Katie Schwartz, FHI 360 (Clinical Research Manager). Study sites and site Investigators of Record: Malawi, Blantyre site (Johns Hopkins University, Queen Elizabeth Hospital): Bonus Makanani; Malawi, Lilongwe site (University of North Carolina, Chapel Hill): Francis Martinson South Africa, Cape Town site (University of Cape Town): Linda-Gail Bekker; South Africa, Durban – Botha's Hill, Chatsworth, Isipingo, Tongaat, Umkomaas, Verulam sites (South African Medical Research Council): Vaneshree Govender, Samantha Siva, Zakir Gaffoor, Logashvari Naidoo, Arendevi Pather, and Nitesha Jeenarain; South Africa, Durban, eThekwini site (Center for the AIDS Programme for Research in South Africa): Gonasagrie Nair South Africa, Johannesburg site (Wits RHI): Thesla Palanee-Phillips Uganda, Kampala site (John Hopkins University, Makerere University): Flavia Matovu Kiweewa Zimbabwe, Chitungwiza, Seke South and Zengeza sites (University of Zimbabwe, University of California San Francisco): Nyaradzo Mgodi Zimbabwe, Harare, Spilhaus site (University of Zimbabwe, University of California San Francisco): Felix Mhlanga. Data management was provided by The Statistical Center for HIV/AIDS Research & Prevention (Fred Hutchinson Cancer Research Center, Seattle, WA) and site laboratory oversight was provided by the Microbicide Trials Network Laboratory Center (Pittsburgh, PA).

Authors Contributions Marla J. Husnik was the lead author and statistical analyst, and Renee Heffron, James P. Hughes, Barbra Richardson, and Ariane van der Straten reviewed, and provided valuable insights into the epidemiological methods used for analysis of the MTN-020/ASPIRE study data. Thesla Palanee-Phillips, Lydia Soto-Torres, Devika Singh, Brenda Gati Mirembe, Edward Livant, Zakir Gaffoor, Leila E. Mansoor, Samantha S. Siva, Sufia Dadabhai, and Flavia Matovu Kiweewa were all involved in the implementation of the MTN-020/ASPIRE study and provided valuable comments/review of the manuscript over several iterations. Finally, Jared M. Baeten was the principal investigator for the MTN-020/ASPIRE study providing leadership to the study team and guidance to the lead author for the writing and analysis of this manuscript.

Funding The MTN-020/ASPIRE study was designed and implemented by the Microbicide Trials Network (MTN). The MTN was funded by the National Institute of Allergy and Infectious Diseases (UM1AI068633, UM1AI068615, UM1AI106707), with co-funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development and the National Institute of Mental Health, all components of the U.S. National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The vaginal rings used in this study were supplied by the International Partnership for Microbicides (IPM).

Declarations

Competing interests JMB is an employee of Gilead Sciences outside of the present work.

References

1. Baeten JM, Donnell D, Ndase P, et al. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. N Engl J Med. 2012;367:399–410.
2. Choopanya K, Martin M, Suntharasamai P, et al. Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet. 2013;381:2083–90.
3. Grant RM, Lama JR, Anderson PL, et al. Pre-exposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med. 2010;363:2587–99.
4. Marrazzo JM, Ramjee G, Richardson BA, et al. Tenofovir based pre-exposure prophylaxis for HIV infection among African women. N Engl J Med. 2015;372:509–18.
5. Van Damme L, Corneli A, Ahmed K, et al. Preexposure prophylaxis for HIV infection among African women. N Engl J Med. 2012;367(5):411–22.
6. Delany-Moretlwe S, Lombard C, Baron D, Bekker LG, Nkala B, Ahmed K, Sebe M, Brumskine W, Nchabeleng M, Palanee-Philips T, Ntshangase J, Sibiya S, Smith E, Panchia R, Myer L, Schwartz JL, Marzinke M, Morris L, Brown ER, Doncel GF, Gray G, Rees H. Tenofovir 1% vaginal gel for prevention of HIV-1 infection in women in South Africa (FACTS-001): a phase 3, randomised, double-blind, placebo-controlled trial. Lancet Infect Dis. 2018;18(11):1241–50.
7. Baeten JM, Palanee-Phillips T, Brown ER, et al. Use of a vaginal ring containing dapivirine for HIV-1 prevention in women. N Engl J Med. 2016;375:2121–32.
8. Hewitt CE, Torgerson DJ, Miles JN. Is there another way to take account of noncompliance in randomized controlled trials? Can Med Assoc J. 2006;175(4):347.
9. Palanee-Phillips T, Schwartz K, Brown ER, et al. Characteristics of women enrolled into a randomized clinical trial of dapivirine vaginal ring for HIV-1 prevention. PLoS ONE. 2015;10(6):e0128857. https://doi.org/10.1371/journal.pone.0128857.
10. Spence P, Nel A, van Niekerk N, et al. Post-use assay of vaginal rings (VRs) as a potential measure of clinical trial adherence. J Pharm Biomed Anal. 2016;125:94–100.
11. Nel A, Haazen W, Nuttall J, et al. A safety and pharmacokinetic trial assessing delivery of dapivirine from a vaginal ring in healthy women. AIDS. 2014;28(10):1479–87.
12. Nel AM, Haazen W, Nuttall J, Romano JP, Mesquita PM, Herold BC, Rosenberg ZF, Niekerk NV. Pharmacokinetics and safety assessment of anti-HIV dapivirine vaginal microbicide rings with multiple dosing. J AIDS Clin Res. 2014;5:355.
13. Nel A, Bekker LG, Bukusi E, et al. Safety, acceptability and adherence of dapivirine vaginal ring in a microbicide clinical trial conducted in multiple countries in sub-Saharan Africa. PLoS ONE. 2016;11: e0147743.
14. Brown ER, Hendrix CW, van der Straten A, et al. Greater dapivirine release from the dapivirine vaginal ring is correlated with

lower risk of HIV-1 acquisition: a secondary analysis from a randomized, placebo-controlled trial. J Int AIDS Soc. 2020;23:11.
15. Van der Weele TJ. Principal stratification—uses and limitations. Int J Biostat. 2011;7(1):28.
16. Frangakis CE, Rubin DB. Principal stratification in causal inference. Biometrics. 2002;58(1):21–9.
17. Shtatland E, Kleinman K, Cain E. Stepwise Methods in Using SAS PROC LOGISTIC and SAS Enterprise Miner for Prediction. In: SUGI'28 Proceedings; 2003.
18. Follmann DA. On the effect of treatment among treatment compliers: an analysis of the Multiple Risk Factor Intervention Trial. J Am Stat Assoc. 2000;95(452):1101–9.
19. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol. 2008;168(6):656–64.
20. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology. 2000;11(5):550–60.
21. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. Biometrics. 2000;56(3):779–88.
22. Hernán MA, Robins JM. Estimating causal effects from epidemiological data. J Epidemiol Community Health. 2006;60(7):578–86.
23. Cole SR, Hernán M. Adjusted survival curves with inverse probability weights. Comput Methods Programs Biomed. 2004;75(1):45–9.
24. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. J Educ Psychol. 1974;56:688–701.
25. Greenland S, Robins JM. Identifiability, exchangeability, and epidemiological confounding. Int J Epidemiol. 1986;15(3):413–9.
26. Husnik MJ, Brown ER, Marzinke M, et al. Implementation of a novel adherence monitoring strategy in a phase III, blinded, placebo-controlled, HIV-1 prevention clinical trial. J Acquir Immune Defic Syndr. 2017;76:330–7.
27. Cole SR, Frangakis CE. The consistency statement in causal inference: a definition or an assumption? Epidemiology. 2009;20(1):3–5.
28. Montgomery ET, van der Straten A, Chitukuta M, et al. Acceptability and use of a dapivirine vaginal ring in a phase III trial. AIDS. 2017;31(8):1159–67.
29. Montgomery ET, van der Straten A, Cheng H, Wegner L, Masenga G, von Mollendorf C, et al. Vaginal ring adherence in sub-Saharan Africa: expulsion, removal, and perfect use. AIDS Behav. 2012;7:1787–98.
30. van der Straten A, Montgomery E, Cheng H, Wegner L, Masenga G, von Mollendorf C, et al. High acceptability of a vaginal ring intended as a microbicide delivery method for HIV prevention in African women. AIDS Behav. 2012;16:1775–86.
31. Griffin JB, Ridgeway K, Montgomery E, Torjesen K, Clark R, Peterson J, et al. Vaginal ring acceptability and related preferences among women in low- and middle-income countries: a systematic review and narrative synthesis. PLoS ONE. 2019;14(11): e0224898.