Facilitating Next-Generation Pre-Exposure Prophylaxis Clinical Trials Using HIV Recent Infection Assays: A Consensus Statement from the Forum HIV Prevention Trial Design Project

Neil Parkin[1], Fei Gao[2], Eduard Grebe[3,4], Amy Cutrell[5], Moupali Das[6], Deborah Donnell[2], Ann Duerr[2], David V. Glidden[4], James P. Hughes[7], Jeffrey Murray[8], Michael N. Robertson[9], Joerg Zinserling[10], Joseph Lau[11], and Veronica Miller[11,*] for the Forum for Collaborative Research Recency Assay Working Group

Standard-of-care HIV pre-exposure prophylaxis (PrEP) is highly efficacious, but uptake of and persistence on a daily oral pill is low in many settings. Evaluation of alternate PrEP products will require innovation to avoid the unpractically large sample sizes in noninferiority trials. We propose estimating HIV incidence in people not on PrEP as an external counterfactual to which on-PrEP incidence in trial subjects can be compared. HIV recent infection testing algorithms (RITAs), such as the limiting antigen avidity assay plus viral load used on specimens from untreated HIV positive people identified during screening, is *one* possible approach. Its feasibility is partly dependent on the sample size needed to ensure adequate power, which is impacted by RITA performance, the number of recent infections identified, the expected efficacy of the intervention, and other factors. Screening sample sizes to support detection of an 80% reduction in incidence for 3 key populations are more modest, and comparable to the number of participants in recent phase III PrEP trials. Sample sizes would be significantly larger in populations with lower incidence, where the false recency rate is higher or if PrEP efficacy is expected to be lower. Our proposed counterfactual approach appears to be feasible, offers high statistical power, and is nearly contemporaneous with the on-PrEP population. It will be important to monitor the performance of this approach during new product development for HIV prevention. If successful, it could be a model for preventive HIV vaccines and prevention of other infectious diseases.

The development of drugs for HIV treatment and HIV pre-exposure prophylaxis (PrEP) has been extraordinarily successful. Safe and well-tolerated fully suppressive antiretroviral therapy (ART) allows people living with HIV to live longer, healthier lives and prevents transmission of HIV to others. Antiretrovirals approved for PrEP can reduce the risk of infection to near zero if taken as prescribed by uninfected persons at risk for HIV and a key component of HIV prevention programs. HIV-incidence is generally declining in pandemic hotspots, such as eastern and southern Africa.[1] Yet, new infections remain unacceptably high in many regions and subpopulations.[2,3]

UNAIDS reports 1.7 million new infections in 2020, primarily among key populations: people who inject drugs, transgender women, sex workers, gay men, other men who have sex with men. These populations and their sexual partners accounted for 93% of new infections outside of sub-Saharan Africa and 35% within sub-Saharan Africa. Sub-Saharan women and girls continue to be at high risk of HIV-infection. Adolescent girls and young women represent 10% of the population yet account for 25% of all new HIV infections in this region.[4,5]

Tenofovir disoproxil fumarate with emtricitabine (TDF/FTC) was approved for prevention by the US Food and Drug Administration (FDA) in 2012, followed by tenofovir alafenamide with emtricitabine (TAF/FTC) in 2019. Long-acting cabotegravir (cabotegravir extended-release injectable suspension) was approved in 2021. Only 1.3 million persons have benefited from these interventions, compared with the 3 million goal UNAIDS set for 2020.[4,5] Challenges in meeting this goal include availability, low uptake, as well as adherence and persistence issues for many.[5–8] In the United States, < 1% of sexually active adults in cities with

[1]Data First Consulting, Sebastopol, California, USA; [2]Fred Hutchinson Cancer Research Center, Seattle, Washington, USA; [3]Vitalant Research Institute, San Francisco, California, USA; [4]Edward Grebe Consulting, Cape Town, South Africa; [5]ViiV Healthcare, Research Triangle Park, North Carolina, USA; [6]Gilead Sciences, Foster City, California, USA; [7]University of Washington, Seattle, Washington, USA; [8]Independent Consultant, Washington, DC, USA; [9]Merck & Co., Inc., Kenilworth, New Jersey, USA; [10]Federal Institute for Drugs and Medical Devices (Bundesinstitut für Arzneimittel und Medizinprodukte, BfArM), Bonn, Germany; [11]Forum for Collaborative Research, Washington, DC, USA. *Correspondence: Veronica Miller (veronicam@berkeley.edu)

Received September 6, 2022; accepted December 8, 2022. doi:10.1002/cpt.2830

high HIV prevalence use PrEP, according to a 2021 Centers for Disease Control and Prevention (CDC) report.[6]

Individuals not receiving PrEP in the aforementioned groups continue to be at risk of HIV with incidence rates exceeding two to three infections per hundred person years. In a recent Johnson & Johnson AdVac platform-based mosaic HIV vaccine phase IIb trial (the Imbokodo trial) HIV incidence was 4.32 per 100 person-years of follow-up in the placebo arm among women enrolled in Malawi, Mozambique, South Africa, Zambia, and Zimbabwe. In 2 trials evaluating the preventive efficacy of neutralizing monoclonal antibodies, placebo incidence rates were 2.98, per 100 person-years of follow-up among men-who-have-sex-with-men (MSM) and transgendered women, and 3.10 per hundred person years in at-risk women in sub-Saharan Africa, respectively.[7] These findings highlight the need for effective, safe, accessible, and implementable HIV prevention options among at-risk populations. Moreover, they emphasize the need for innovation in HIV prevention research to ensure that the standard-of-care is provided for all participants while allowing determination of efficacy, referred to as assay sensitivity of the trial.

Developing new HIV prevention options that can be delivered effectively globally will require innovation in trial design and collaboration across stakeholders. Support from the HIV-affected community at the global level is essential for new products to be tested in at-risk communities and accepted by policy makers for use.

The Forum for Collaborative Research (the Forum), a public-private partnership, brings together all stakeholders in a neutral and independent venue for open dialogue and deliberation to address issues in areas of unmet medical need in different disease areas, including HIV.[9,10] HIV Forum members requested the Forum to facilitate consensus for innovative clinical trial designs to accelerate access to effective and safe PrEP products meeting the needs of all at-risk individuals. Individuals participating in this project bring expertise from patients and at-risk community organizations, regulatory agencies, the World Health Organization (WHO), regulated industry, and HIV prevention clinical science and research.

Members reviewed the statistical and logistic challenges presented by traditional (superiority or noninferiority designs), reviewed elsewhere,[11–13] and explored the suitability of an external counterfactual approach. Traditional placebo arms represent the most accurate counterfactual estimate[14]; if a randomized placebo arm is not available, external counterfactual estimate of HIV-incidence among people not on PrEP could be derived using several approaches. For example, the use of rectal gonorrhea incidence as a predictor of HIV incidence in MSM not on PrEP was suggested as one possibility earlier on.[15] A subsequent analysis to evaluate whether the correlation between rectal gonorrhea and HIV incidence reported in ref. 15 might be useful to predict HIV incidence in a large MSM cohort that did not support this approach for MSM PrEP trials.[16] Ongoing analyses are evaluating the usefulness of past-trial placebo cohorts (Miller, personal communication). An approach warranting further investigation is the adherence-efficacy relationship of emtricitabine and tenofovir fumarate to calculate HIV incidence in trials with this specific control arm.[17] Other information sources might include surveillance and/or epidemiological studies.[18] In particular, cross-sectional incidence studies based on HIV recency assays[19] are of interest.

Following several rounds of discussion, the Recency Assay Working Group was established and tasked with reviewing the feasibility of HIV recency assays as one method to construct a counterfactual (without use of PrEP) HIV incidence estimate for use as an external control in prevention trials. The review and discussion included assay type, statistical methods for sample size calculations for different epidemiological contexts, and general implementation strategies in studies of new PrEP products in different populations. This manuscript is the result of this review and discussion.

DRUG DEVELOPMENT AND REGULATORY CHALLENGES FOR SECOND GENERATION PREP PRODUCTS

The development of new PrEP products is challenged by the high efficacy and effectiveness of approved and tested interventions.[18,20–25] As HIV seroconversion event rates decrease on approved PrEP regimens, sample sizes become prohibitively large for active-controlled trials assessing superiority or noninferiority.

The traditional pathway for approval of new drugs requires adequate and well-controlled clinical trials,[26] either placebo-controlled trials, or standard of care (SOC) controlled trials with superiority or noninferiority designs. Comparison against a placebo or no treatment arm is no longer ethical in HIV prevention[27] limiting traditional design options to superiority or noninferiority comparisons to the SOC. Such randomized trials require large sample sizes to achieve sufficient end points for statistical comparison. Whereas superiority trials may be feasible if the new agent has substantial advantages vs. SOC (e.g., improved adherence with long-acting PrEP agents), a noninferiority design is not an ideal option due to inconsistent treatment effects of some approved regimens across subpopulations and due to need for large sample sizes. For MSM and transgender women, the constancy assumption based on the original placebo-controlled trial (demonstrating low efficacy)[21] is no longer valid given the much higher efficacy observed with improved adherence[25,28]; for cisgender women, demonstration of efficacy vs. placebo was inconsistent and/or negligible in previous trials of oral drugs, precluding the estimation of reliable noninferiority margins.

In this article, we review a feasible path for the approval of new PrEP products using a counterfactual estimate based external control and describe how an HIV recency assay can contribute to derive such an estimate.

USE OF EXTERNAL CONTROLS AND REGULATORY GUIDANCE

Regulatory guidance provides for use of concurrent or historic external controls for situations in which a traditionally controlled trials are not ethical and/or feasible.[29] The International Council on Harmonization (ICH) E10 guidance discusses limitations and approaches for externally controlled trials. Control patients should be as similar as possible to the population receiving the study group and should be selected before performing comparative analyses. If no single optimal external control exists, then multiple external controls should be used to draw inferences. The study group outcomes should be substantially superior to the most favorable control to conclude efficacy. A pertinent example of the use of external controls in clinical trials supporting regulatory approval is the use of a historical reference as a control

in hormonal contraceptives trials. Many hormonal contraceptives have been approved in the United States and other countries using this approach.[30] For PrEP trials, this entails building an external control to estimate what the HIV incidence would have been in the absence of PrEP (a counterfactual placebo estimate) to be able to conclude with confidence that an observed low incidence rate could be attributed to an investigational drug's preventive efficacy, or whether the trial did not enroll sufficient numbers of high-risk individuals from those communities.

Ideally, the counterfactual HIV incidence estimate is informed by multiple data sources, including ongoing surveillance, epidemiologic studies, and past trials performed in the same regions. Because of changing HIV demographics, a counterfactual estimate that is based on concurrent information is preferred.

DIFFERENTIATING BETWEEN RECENT AND CHRONIC HIV INFECTION

Estimating cross-sectional HIV incidence requires a way to capture incident infections. In the absence of community-wide HIV incidence cohort studies, inferring HIV incidence through estimation of recent infections is needed.

In this section, we review the status of knowledge of differentiating between recent and chronic HIV infection and discuss how this approach could be used in PrEP trials, from a trial sponsor, clinical researcher, and regulatory perspective. Note that none of the tests have been approved by the FDA for differentiation of recent and chronic infection; recency testing has been performed for research and surveillance purposes only.

Most laboratory tests that distinguish between recent and long-standing HIV-1 infection rely on the well-characterized evolution of the humoral antibody (Ab) response to HIV infection and are reviewed elsewhere.[31–39]

One important challenge to the interpretation of tests for recent infection is the inherent biological variability in the kinetics of Ab maturation during infection between individuals; this can be influenced by particular host, viral, and external characteristics, including sex, pregnancy status, HIV-1 subtype, ART status, etc. Therefore, recent infection testing performed for the purpose of estimating incidence in a population requires a probabilistic approach, based on the test performance described by two key parameters: the mean duration of recent infection (MDRI) and false recency ratio (FRR; sometimes referred to as the false recent rate, or proportion false recent).[40–42] The MDRI is defined as the average number of days spent since the date of earliest detectable infection while still being classified as "recent" by measurement of a particular biomarker (usually, Ab avidity and/or concentration). The FRR is the percentage of "recent" test results obtained from individuals who are known to have been infected for a period longer than an arbitrarily assigned "recent" time cutoff (known as T), usually 1 to 2 years.[41,43] MDRI and FRR are established for specific contexts through assay calibration studies using well-characterized, dated specimens from HIV-infected individuals with documented diagnostic test histories that support estimates of infection dates.[44]

For estimation of HIV incidence with reasonable precision, the MDRI and FRR must be within certain ranges. Target MDRI and FRR values for different use cases have been delineated elsewhere.[45] Unacceptably high FRRs occur in groups of individuals with long-standing infection but who are elite controllers, who have suppressed viral load (VL) on ART, or are infected with subtype D HIV-1.[46–48] High FRR in patients on ART is most likely a result of "reversion" of the HIV-1 Ab response, such that it resembles that seen in recently infected people (i.e., relatively low avidity and/or concentration). To reduce the FRR in these populations, additional tests can be added in combination with the recent infection test to form a recent infection testing algorithm (RITA). For example, VL testing can identify individuals with a false positive recent infection result and low VL; such individuals are classified as not being recently infected, based on the assumption that recently infected individuals with very low VL (e.g., below 1,000 copies/mL) are exceedingly rare. Similarly, if the MDRI is too low, a prohibitively large sample size will likely be required to identify enough recent infections to support a precise incidence estimate, because the average time spent in the recent infection window is short.

Different laboratory tests for recent HIV-1 infection have been described, and their performance evaluated (as a single test or as part of a RITA) using well-characterized panels of specimens.[47–55] The selection of the most appropriate assay or algorithm for a specific application depends on the desired level of precision of assay-based incidence estimates, as well as many other practical issues associated with implementation.

TYPES OF RECENCY TESTS

The most widely used and best characterized research recency assays is the limiting antigen antibody avidity assay (LAg).[51] Originally developed at the US CDC, it relies on detection of low avidity Abs that bind to an immunodominant region of the gp41 envelope protein. The LAg assay is available commercially for research purposes only and has been extensively characterized both on its own or in a RITA in combination with VL in assay calibration studies[47,48] and validated in the field.[56] The LAg + VL RITA has MDRI and FRR that meet the target product profile for incidence estimation in population surveys,[45] and had been widely used for large-scale surveillance efforts in many settings.[19,57–62] An External Quality Assurance proficiency testing program for the LAg assay has been established (https://eqapol.dhvi.duke.edu/programs/lag)[63]; participation by all laboratory tests implementing this test is strongly recommended.

An alternative to LAg that also meets the criteria listed above is the ARCHITECT HIV Ag/Ab Combo Assay.[64] The approved intended use of this test is as a qualitative aid to diagnosis of HIV infection; however, the quantitative information contained in the signal to cutoff ratio has been used in research to discriminate recent from long-term infections.[64–66] Because this test relies in part on anti-HIV antibody levels, it is also subject to false classification of infection as recent when viral replication is suppressed and should be used in combination with VL tests in a RITA to reduce the FRR. Relatively less is known about the optimal threshold to use when implementing a RITA using ARCHITECT, compared to the LAg assay. Note that the discrimination of recent from long-term infections is an off-label use for which the test has not been approved by the FDA.

USING HIV RECENCY TESTING TO ESTIMATE COUNTERFACTUAL HIV INCIDENCE FOR PREP TRIALS

The use of HIV recency assays to derive a counterfactual incidence estimate in the context of clinical trials for new PrEP interventions is illustrated conceptually in **Figure 1**. During the screening process, participants who are HIV-negative and meet other trial inclusion criteria are enrolled in the study and receive the intervention. Incidence in this population can be derived by direct observation, with new infections being identified through HIV testing at regular intervals (e.g., monthly). Specimens from potential participants who are found to be HIV-positive at screening are subjected to recency testing. An incidence rate can be calculated using MDRI and FRR values that correspond as closely as possible to the subtype of the infection and to the method used for identifying the HIV-positivity (see below). This estimate of the rate of infection in the population meeting the risk profile without use of the trial product(s) is the "placebo counterfactual incidence rate." The precision of the placebo counterfactual incidence estimate will depend on the number of participants screened (negative and recent) and the number of recently infected cases identified, as well as uncertainty associated with the MDRI and FRR.

The LAg assay (with a normalized optical density threshold of 1.5) in combination with VL is the best characterized and most implemented test (**Figure 2**). Alternatives to LAg exist or may be developed in the future. Although less data to support robust MDRI and FRR estimates exist for the alternatives, they may have advantages and their utility should be investigated. The choice of assay(s) and the thresholds used to classify participants as recently infected, should be established before the trial is initiated.

The LAg avidity assay is commercially available for research purposes from two manufacturers: Sedia Biosciences (Beaverton, OR) and Maxim Biomedical (Rockville, MD). These kits can use plasma or dried blood spot specimens. In a comparative evaluation performed by the Consortium for Evaluation of Performance of HIV Incidence Assays (CEPHIA), MDRIs differed between manufacturers by about 32 days.[67] Although well-correlated with each other, normalized optical density (ODn) values from the Maxim assay were lower than with the Sedia assay, because of a difference in the calibrator control sample included in the kits.

The standard threshold for ODn that is used to classify an individual as having been recently infected is 1.5.[50] Alternative thresholds are possible and may support more precise incidence estimates in specific contexts. For example, raising the threshold to 2.0 would be expected to increase the MDRI, because more individuals would be classified as being recently infected. However, the FRR would also be expected to be higher, and the optimal balance between these two parameters will depend on differences in the target population(s) such as ART coverage, subtype distribution, and the actual incidence and prevalence.[66]

The threshold applied to the VL result depends on the specimen type collected at screening. For dried blood spot, a VL threshold of 1,000 copies/mL should be applied, because the sensitivity of VL testing in DBS is lower than in plasma (higher limit of detection). For plasma, a lower threshold can be used, and is advantageous because it lengthens the MDRI, leading to increased ability to detect recent infections and potentially resulting in improved precision of the incidence estimate. MDRI and FRR for a RITA consisting of the Sedia LAg assay plus VL for four major subtypes, derived from analysis of algorithm calibration experiments from individuals who are HIV infected and ART-naïve in the CEPHIA repository,[67] are shown in **Table 1**. MDRI and FRR for a RITA consisting of the Abbott ARCHITECT assay plus VL (using a threshold of 75 copies/mL) for 4 major subtypes, derived from the same calibration data set,[66] are shown in **Table 2**.

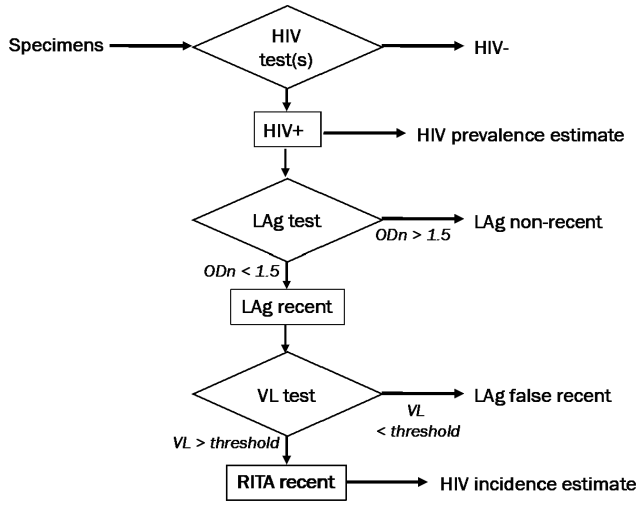

Figure 1 Overview of clinical trials of PrEP including recent infection testing to estimate the counterfactual incidence. FRR, false recent ratio; LAg, limiting antigen antibody avidity assay; MDRI, mean duration of recent infection; PrEP, pre-exposure prophylaxis; RITA, recent infection testing algorithm; VL, viral load.

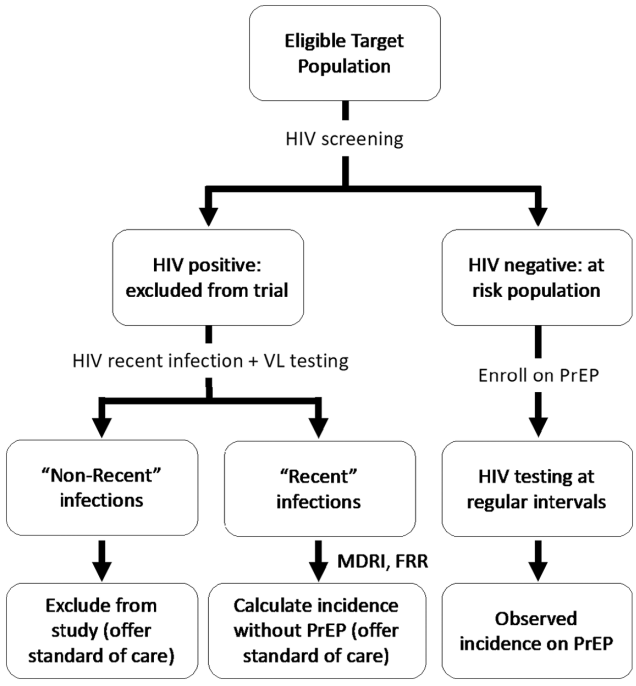

Figure 2 A Recent Infection Testing Algorithm. FRR, false recent ratio; LAg, limiting antigen avidity assay; ODn, normalized optical density; RITA, recent infection testing algorithm; VL, viral load.

KNOWLEDGE GAPS AND CAVEATS FOR USE OF RITAS

HIV testing frequency and early ART

Assay and algorithm calibration information is most reliably applied for incidence estimation when the characteristics of the population of trial participants are well-matched to those of the population used to derive the calibration parameters (i.e., MDRI and FRR). Nearly all available assay calibration data were generated many years ago, when HIV testing was less frequent as it is likely to be currently, and when early ART was not yet the SOC. In the case of clinical trials of new PrEP agents, the distribution of times since infection among HIV-infected individuals sampled could be different if the target population is one in which many people are frequently tested for HIV (e.g., every 3–12 months), and placed on ART immediately if diagnosed, because they would not present for screening. In a population where testing is frequent and early treatment is common, individuals "exit" the recent state by getting diagnosed and treated; ART lowers VL sufficiently that they would be classified as not recently infected (see **Figure 2**). In such a population, the MDRI of the RITA would be shorter than that based on the population from which calibration data are derived. Application of an MDRI that is too high would lead to an incidence estimate whose value is lower, and whose precision is higher, than in reality. In such a case, the efficacy of a new PrEP treatment would be evaluated against a lower placebo incidence rate, thus biasing against the new treatment. This effect could be further complicated by heterogeneity in frequency of testing and early ART initiation in specific subgroups of people. Careful consideration should be given to the specifics of each context, especially testing frequency and the proportion of patients who are HIV-positive and on ART. In some situations, an adjustment to MDRI and/or FRR (or other parameters, such as T) may be needed. Such adjustments, the details of which have not yet been described, could affect the utility of calibration data shown in **Tables 1** and **2**.

HIV-1 subtype diversity

The HIV epidemic is characterized by a high degree of viral genetic variability with important implications for transmission, diagnosis, treatment, and prevention.[68] Of three genetic groups (M for major, O for outlier, and N for non-M and non-O, group M is responsible for most infections globally, and comprises nine subtypes (A-D, F-H, J, and K)). Furthermore, circulating recombinant forms (CRFs) are common. Their relative distributions of group M subtypes and CRFs are reviewed in ref. 68. Because of a paucity of specimens that belong to subtypes other than those shown in **Tables 1** and **2** in the panels used to calibrate the algorithm (i.e., determine the MDRI and FRR), surveys conducted in some regions of the world where such subtypes are prevalent will need to use an MDRI and FRR derived from much smaller data sets (e.g., CRF01_AE)[50] or estimated in the absence of complete calibration data (e.g., subtype F, G, CRF02_AG, or other CRFs).[69] For CRF01_AE, the published MDRI for the LAg assay was derived using slightly different methods than those shown in **Table 1**, and thus may not be directly comparable to these values; furthermore, an FRR for CRF01_AE was not reported.[50] However, the MDRI that has been calculated is very similar to the subtype B estimate (122 vs. 129 days), suggesting that using subtype B MDRI and FRR for CRF01_AE may be an acceptable approximation.

The FRR for subtype D is significantly higher than for other subtypes; this has been reported previously for LAg and other serological recency tests.[70–73] The precision of incidence estimates in countries where a significant proportion of subtype D infections are expected, such as Chad, Kenya, South Sudan, Tanzania, and Uganda, may suffer as a result.

Trials that involve populations in multiple areas, each of which has different prevalence of various subtypes, should employ weighted MDRI and FRR values based on the relative proportions of each subtype, determined experimentally from sequence data, or estimated based on previously gathered molecular epidemiological data.

ARV drug level testing

Some investigators have included testing for ARV drug levels in a RITA based on LAg + VL. The rationale for ARV testing is

Table 1 MDRI and FRR for LAg (Sedia) and VL-based recency testing algorithm

HIV-1 subtype	VL >	MDRI[a] (days) (95% CI)	Untreated FRR (%) (95% CI)
All	75	202 (180–224)	1.7% (0.4–4.9)
	1,000	171 (152–191)	1.1% (0.1–4)
A	75	212 (158–274)	2.6% (0.1–13.8)
	1,000	186 (137–245)	2.6% (0.1–13.8)
B	75	189 (145–239)	1.8% (0.1–9.6)
	1,000	176 (132–226)	0% (0–6.4)
C	75	194 (169–222)	1.4% (0–7.6)
	1,000	162 (141–185)	1.4% (0–7.6)
D	75	262 (168–375)	NA
	1,000	209 (126–307)	NA

ART, anti-retroviral treatment; CI, confidence interval; FRR, false recency ratio; LAg, limiting antigen antibody avidity assay; MDRI, mean duration of recent infection; NA, not applicable; ODn, normalized optical density; VL, viral load.

[a]MDRI based on LAg ODn < 1.5, T = 2 years, and HIV infection detection using a hypothetical test with a sensitivity of 1 copy/mL. When using fourth generation Ag/Ab tests, the MDRI should be shortened by 11 days. FRR shown is for untreated patients. FRR for treated patients is 0% for all subtypes (95% CI 0–2.8%). A weighted FRR based on the estimated proportion of patients on ART should be used. NA: not available, ≤ 10 individuals infected with subtype D for ≥ 2 years.

Table 2 MDRI and FRR for ARCHITECT and VL-based recency testing algorithm

HIV-1 subtype	VL>	S/CO<	MDRI[a] (days) (95% CI)	FRR (%) (95% CI)
All	75	150	154 (135–175)	2.2% (0.6–5.7)
	75	200	192 (169–215)	5.1% (2.3–9.4)
	75	250	234 (208–261)	7.3% (3.9–12.2)
A	75	150	156 (121–195)	5.3% (0.6–17.7)
	75	200	224 (167–288)	7.9% (1.7–21.4)
	75	250	262 (199–335)	10.5% (2.9–24.8)
B	75	150	145 (103–191)	0% (0–6.4)
	75	200	169 (121–220)	1.8% (0–9.6)
	75	250	211 (152–275)	1.8% (0–9.6)
C	75	150	139 (116–163)	0% (0–5.1)
	75	200	169 (141–200)	4.2% (0.9–11.9)
	75	250	210 (177–246)	5.6% (1.6–13.8)
D	75	150	241 (145–347)	NA
	75	200	296 (189–406)	NA
	75	250	339 (231–456)	NA

ART, anti-retroviral treatment; CI, confidence interval; FRR, false recency ratio; MDRI, mean duration of recent infection; NA, not applicable; ODn, normalized optical density; S/CO, signal to cutoff ratio; VL, viral load.
[a]MDRI based on $T = 2$ years, HIV infection detection using a hypothetical test with a sensitivity of 1 copy/mL. When using fourth generation Ag/Ab tests, the MDRI should be shortened by 11 days. FRR shown is for untreated patients. FRR for treated patients is 0% for all subtypes (95% CI 0–2.8%). A weighted FRR based on the estimated proportion of patients on ART should be used. NA, not available, ≤10 individuals infected with subtype D for ≥2 years.

that some individuals with longstanding HIV infection who are on ART can have LAg ODn < 1.5 (because of prior "reversion" of the immune response) and VL > 1,000 copies/mL because of incomplete adherence or the resurgence of drug-resistant virus. These individuals would be incorrectly classified as having recent infection by the standard LAg + VL RITA. This three-test RITA has been implemented in large surveys in African countries, using an MDRI of 130 days (likely to be an overestimate) and assumed FRR of 0%.[58] However, this approach is complicated by the lack of calibration information, because ARV testing has not been performed on the large specimen panels used to date for derivation of MDRI and FRR, and there is a paucity of specimens from patients not virologically suppressed on ART in these panels. It is expected that the MDRI and FRR would both be lower. Furthermore, kinetics of anti-HIV Ab avidity changes in patients with viral replication ongoing in the presence of ART are poorly characterized. A recent report describing the implementation of a LAg-based RITA in resource-limited settings included only 3 of 92 patients being re-classified as non-recent following ART drug level testing.[74] The body of supporting information, in addition to additional cost and logistical complexity, does not support including ARV testing in a RITA for the purpose of generating a counterfactual incidence estimate for clinical trials of PrEP.

SAMPLE SIZE CALCULATION

One of the main determinants of the feasibility of this approach is the total screening sample size (not just the number of participants actually enrolled in the clinical trial) required to provide sufficient precision around the laboratory-based incidence estimate, because the number of people to be screened and the number enrolled in the trial on PrEP are major drivers of the cost associated with clinical trial implementation. Important variables that impact sample size are outlined in **Table 3**.

Building on previous work using the *inctools* R package,[75] total screening sample size is calculated based on a test statistic of log of the ratio of experimental arm and placebo arm incidences,[76] where the placebo arm incidence is estimated by estimator in Kassanjee *et al*.[41] Specifically, asymptotic distributions of the log estimated incidences are combined to derive the asymptotic distribution of the test statistic under null and alternative hypotheses, allowing for calculation of the sample size with desired type-1 error and power under specific null and alternative settings of intervention efficacy.

To illustrate how some of the variables listed in **Table 3** impact sample sizes, we present 4 hypothetical scenarios in **Table 4**, using incidence and prevalence data for several key populations with high incidence. These scenarios are not intended to precisely represent the ranges of incidence and prevalence in populations likely to be considered for clinical trials of PrEP, but serve as examples of key populations where PrEP is urgently needed. The sample sizes calculated are for a single investigational PrEP arm and comparison to the RITA-derived incidence estimate. Based on comparison to the numbers of participants in recently completed or ongoing trials of various PrEP agents,[18,21–24,80–83] the required sample sizes shown in **Table 4** are within a range that would be likely feasible for a large study sponsor. Preliminary investigation into sample sizes needed for trials that also include an active control comparator arm indicate that these trials are also feasible, depending on the level of precision stipulated for the comparison between arms (J.P. Hughes *et al*., unpublished data). Other clinical trial designs could necessitate larger sample sizes; methods for determining sample sizes for different designs are in development.

The scenarios modeled above show that the proposed approach is not limited by large sample size requirements. To extend this

Table 3 Key variables that determine required sample sizes

Variable	Rationale	Comments
Desired level of precision	Higher precision requires larger sample sizes which will allow greater opportunity to observe events (i.e., recency assay positive infections and new infections on PrEP)	The level of precision that is required to support the determination of a statistically significant difference in the two incidence estimates can be determined based on accepted levels of type-1 error and power
Expected prevalence and incidence	For maximum precision and smallest possible sample size, the proportion of HIV-positive cases that meet the recency assay/algorithm criterion should be as high as possible, and the background of cases of long-standing infection as low as possible	High prevalence leads to a reduction in the number of individuals at risk for HIV acquisition, which is a component of the denominator in the incidence rate calculation
Expected efficacy of the intervention	An intervention that is expected to have a very large effect, such as ≥90% reduction in incidence compared to the counterfactual estimate, will require less precision in order to reach statistical significance of the reduction	Minimum acceptable effect of the intervention under the null hypothesis, i.e., the intervention effect we would like to rule out, also impacts the sample size
Assay calibration (MDRI and FRR)	Longer MDRI and lower FRR will permit the accurate detection of more recent infections compared to shorter MDRI or higher FRR	Calibration parameters are specific to the assay or RITA and the target population (e.g., based on HIV-1 subtype prevalence)
Trial design (e.g., number of different treatments being evaluated)	Evaluation of more than one intervention, or inclusion of an active control arm, will significantly impact required sample sizes	

observation, the impact of MDRI, FRR, incidence, and prevalence on sample sizes is illustrated in **Figure 3**. For a given incidence and prevalence, increasing the MDRI from 120 to 200 days, keeping other assumptions constant, leads to a decrease in sample size, especially when FRR is 3% or higher. Similarly, sample sizes are lower for lower FRR values, with the magnitude of the effect being minimized when MDRI is high (e.g., over 180 days; **Figure 3a**). When MDRI is low (e.g., below 150 days) and FRR is 4% or higher, the sample sizes become so high that the trial may not be feasible. For a specific combination of MDRI and FRR, samples sizes increase with lower incidence or higher prevalence; when incidence is over ~5% per year, the impact of higher prevalence is lessened. At low prevalence (e.g., 10%), sample sizes remain feasible at lower values of incidence (**Figure 3b**).

CLINICAL TRIAL DESIGN CONSIDERATIONS

We assume that the primary endpoint for new PrEP product trials will be reduction of HIV incidence in the new PrEP product arm

Table 4 Sample sizes required for clinical trials using RITA-based counterfactual incidence estimates

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Location(s)	KwaZulu-Natal, South Africa[77]	Mpumalanga province, South Africa[78]	South Africa, eSwatini, Kenya, and Zambia[79]	Peru[7]
Population	AGYW 14–17 years old	MSM > 18 years old	AGYW 16–35 years old	MSM > 18 years old
Survey period	2004–2007	2012–2015	2016–2018	2016–2018
Incidence (annual)	4.7%	12.5%	3.8%	6.1%[b]
Prevalence	27.6%	32.4%	12.1%[a]	29.0%[b]
Subtype	C	C	C	B
MDRI (days)	151	151	151	165
FRR[c]	1.4%	1.4%	1.4%	0%
Number on PrEP[d]	1,369	454	1,469	904
Total number to be screened	2,101	747	1,857	1,414

AGYW, adolescent girls and young women; ART, anti-retroviral treatment; FRR, false recency ratio; MSM, men having sex with men; PrEP, pre-exposure prophylaxis.

[a]Prevalence based on number of HIV positive individuals identified during screening for the ECHO trial.[79] [b]Incidence from placebo group from the AMP study in Lima[7]; prevalence from surveillance study of HIV positive MSM and transwomen; Ministry of Health of Peru, Coordinadora Multisectorial Multisectorial en Salud (CONAMUSA) and Jorge Sanchez, personal communication. [c]FRR derived from ART naïve patient population. If the proportion of people infected with HIV and on ART can be estimated, a weighted FRR (assuming no false recent results in treated patients) should be used instead. [d]The number enrolled into the PrEP intervention arm(s), assuming a single arm in the trial. All sample size calculations performed assuming: fourth generation Ab/Ag testing to identify HIV-positives, 80% reduction (null hypothesis: 50% reduction) in incidence on PrEP, 90% of individuals who are HIV-negative enroll on PrEP, 90% of HIV positive specimens yield valid recency testing results, two years of follow-up on PrEP, 7.4% and 14.5% relative standard error on the MDRI for subtype C and subtype B, respectively, 25% RSE on FRR, significance level 0.05 and power 0.8. https://github.com/feigao1/samplesize_RA/.

compared with the counterfactual HIV incidence estimate derived using the recency assay and confirmed using other regional data sources. Use of an active PrEP SOC control arm for additional secondary efficacy and safety comparisons of the new PrEP agent vs. the active PrEP SOC control arm could also be employed to strengthen the trial.

The concept of calculating a counterfactual HIV incidence for the purpose of evaluating the efficacy of HIV prevention products is novel; operationalizing it involves several challenges which must be considered. In order for the statistical framework described above to be valid (i.e., a reduced incidence observed on PrEP compared with the incidence estimate is truly a result of the intervention), the following assumptions have been made.

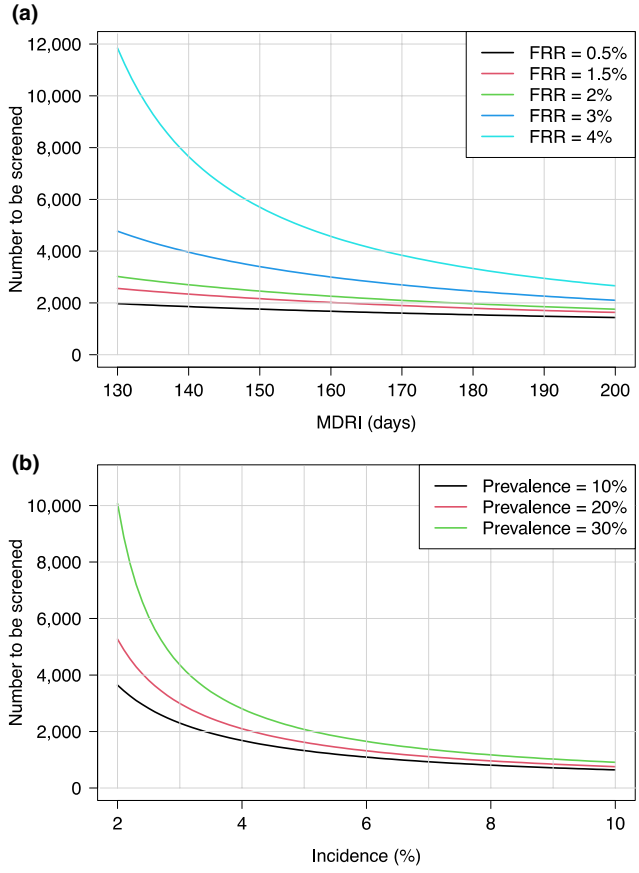

Figure 3 Impact of incidence, prevalence, MDRI and FRR on required sample sizes. Sample sizes shown are the number to be screened, assuming a single arm in the trial. All sample size calculations performed assuming: fourth generation Ab/Ag testing to identify HIV-positives, 80% reduction (null hypothesis: 50% reduction) in incidence on PrEP, 90% of individuals who are HIV-negative enroll on PrEP, 90% of HIV positive specimens yield valid recency testing results, two years of follow-up on PrEP, 14.5% relative standard error on the MDRI, 25% RSE on FRR, significance level 0.05 and power 0.8. https://github.com/feigao1/samplesize_RA/. (**a**) Effect of FRR on sample size over a range of MDRI values (assumes a baseline incidence of 4.7% and prevalence of 27.6%). (**b**) Effect of baseline prevalence on sample size over a range of incidence (assumes MDRI of 151 days and FRR of 1.4%). FRR, false recent ratio; MDRI, mean duration of recent infection; PrEP, pre-exposure prophylaxis; VL, viral load.

1. HIV acquisition risk factors in people who test HIV-positive at screening and those eligible and enrolled in the trial are well-matched. In order for the comparison between incidence on PrEP and the counterfactual incidence derived from recent infection testing to be valid (i.e., that the expected reduced incidence observed on PrEP is truly a result of the intervention), the HIV risk acquisition factors must be well-matched in the two groups. This could most readily be achieved by applying eligibility criteria *before* an HIV test is performed. Doing so represents a significant change compared with how clinics usually identify eligible trial participants. Most clinical sites, when evaluating participants for inclusion in an HIV prevention trial or for clinical PrEP, would first test the participant for HIV, and refer them to a clinical prevention trial or PrEP intervention if negative, or to clinical HIV care and treatment if positive. Thus, implementation of the screening cohort will require trial sites to ensure that all clinic visitors who meet criteria for eligibility for the randomized cohort, are indeed, screened for HIV, so that those who are positive can have a recency test done and contribute to the background HIV incidence calculation. Pre-screening by other providers in the area or by the potential participant themselves so that only individuals known to be HIV negative present for screening this will bias the HIV incidence estimate in the screening cohort to be lower than the true incidence. Similarly, efforts should be made to ensure that the population of HIV-negative screened participants is similar to those who are eligible for and willing to use PrEP.
2. Willingness to be tested for HIV as part of the screening process is independent of HIV status. If people who have recently been diagnosed with HIV infection (e.g., within the last year) do not present for screening, then the number of recent infections counted and the resulting incidence estimate will be too low. It is not possible to require screening participants to reduce the frequency of testing in order to participate in new PrEP trial research, as quarterly testing is recommended for persons at risk.[84]

EXTENSION OF SCREENING PERIOD TO INCREASE INCIDENCE ESTIMATE PRECISION

To ensure that the precision of the counterfactual incidence estimate is sufficient, if the number of RITA-positive recent infections are identified is lower than expected, sample collection for the HIV incidence cohort may need to continue beyond the time at which the target numbers of people on PrEP have been reached. However, this raises additional questions about the responsibility of the sponsor to provide treatment and care for individuals with HIV infection identified during this period, or to provide PrEP to individuals who are not infected and could not be enrolled in the clinical trial component of the study.

POTENTIAL FOR TRIAL INCLUSION OF INDIVIDUALS ALREADY ON PREP

In populations at high risk of acquiring HIV, where new PrEP trials are likely to be conducted, there may be individuals who are already using approved interventions (e.g., TDF/FTC) but who wish to participate in the study of alternatives. This has implications for

the design of a trial that relies on a counterfactual placebo estimate, because the risk of HIV acquisition in people on PrEP is expected to be lower than the theoretical placebo group. Furthermore, if such individuals are found to be HIV-infected at screening (e.g., due to incomplete adherence), it is not clear whether the kinetics of antibody maturation would be the same as in individuals not on PrEP, although it is likely that this is delayed to some extent.[85,86] For these reasons, while not preventing these individuals from participating in the trial, they should be excluded from the cohort used to generate the counterfactual incidence estimate. Alternatively, a sensitivity analysis can be conducted, wherein an incidence is also calculated regardless of history of prior PrEP use.

CONCLUSIONS

The use of HIV recent infection testing to generate a counterfactual incidence estimate is a promising new approach that can facilitate implementation of clinical trials for new prevention interventions. If integrated into the screening process of the clinical trial, recent infection testing could address the goal of providing a background incidence estimate that is more contemporaneous and similar with respect to the baseline characteristics of participants in the new prevention intervention arm than those generated by other methods. Recent infection testing could serve as one of several methods to establish a counterfactual placebo estimate of HIV incidence in specific communities.

This approach is new, relies on assumptions, and has several important challenges that need to be addressed. However, a high prevention efficacy, resulting in a very low incidence of HIV infection on PrEP, will allow for a clear difference with the counterfactual estimate, and compensate for concerns about the precision of the estimate. It will be important to carefully evaluate the proposed approach as new prevention studies are rolled out. Open and ongoing communication between sponsors, communities, clinical researchers, and health authorities will be essential to support the acceptability of this method. As recency data are generated and estimates confirmed with other data sources, cross-stakeholder collaborations will be critical to monitor progress using our proposed approach for assessing efficacy, and its impact on communities participating in such trials.

Based on the current status of knowledge, as outlined in this review, the Forum for Collaborative Research and the Recency Assay Working Group propose the following:

1. The use of a recent infection assay and algorithm (a combination of a recency assay and other assays such as HIV viral load) should be considered as one method to generate a counterfactual incidence estimate, by testing specimens from individuals infected with HIV-1 from the same population as PrEP trial participants.
2. When using a recency-based counterfactual estimate in a primary efficacy comparison, other incidence estimate methodologies should be used to triangulate/corroborate the recency estimate. Other methodologies could include: epidemiologic estimates in the area in which the trial will be conducted, use of previous placebo estimates at the same trial site, and post-randomization incidence data methods (e.g., adherence-efficacy or STI/HIV incidence methods).[15,17,87]
3. The characteristics of the population used for the assay calibration should be well-matched to that being targeted for prospective testing to estimate incidence.
4. Specifically, for use in generation of a counterfactual incidence estimate in the context of clinical trials of PrEP, the following criteria should be applied when selecting a recency assay or algorithm:
   - MDRI and FRR values should have been determined, *a priori*, and be within ranges that can support incidence estimation with the required level of precision.
   - High quality kits should be available from one or more manufacturers in a sustainable fashion.
   - Quality assurance for test operation should be implemented in each participating laboratory (validation and external QA).
   - Minimal, and preferably no modifications to validated assay procedures should be required, assuming multi-laboratory implementation, because standardization is needed.
   - Cost should be affordable for trial sponsors.
   - Specimen types needed should not differ from those typically collected in clinical trials.

ACKNOWLEDGMENTS

The authors thank Charu Mullick and Thamban Valappil from the US Food and Drug Administration for their insightful comments and review of this manuscript. We thank Tamar Tchelidze (Forum for Collaborative Research) for organizational support of the Recency Assay Working Group.

The Forum for Collaborative Research HIV Prevention Trial Design Project members include:

The Forum for Collaborative Research Recency Assay Working Group has the following members in addition to all co-authors: Yen Pottinger (Columbia University); Regine Lehnert (Federal Institute for Drugs and Medical Device); Timothy D. Mastro (FHI360); Beatriz Grinsztejn (FIOCRUZ); Tamar Tchelidze (Forum for Collaborative Research); Jared Baeten, Christoph Carter, Stephanie Cox, Ramin Ebrahimi, Alex Kintu, James Rooney, and Yongwu Shao (Gilead Sciences); Jessica E. Justman (ICAP at Columbia; Mailman School of Public Health); Zeda Rosenberg (International Partnership for Microbicides); Oliver Laeyendecker and Susan H. Eshelman (Johns Hopkins School of Medicine); Frances Cowan (Liverpool School of Tropical Medicine); Brian Rice (London School of Hygiene & Tropical Medicine); Joe Ma (Maxim Biomedical); Elizabeth Russell and Kathleen Squires (Merck); Ronald Mink (Sedia Biosciences); Alex Welte (Stellenbosch University); Kimberly Struble and Wen Zeng (US Food and Drug Administration); David Dunn (University College London); Peter J. Dailey and Sandra McCoy (University of California; Berkeley); George Rutherford and Susie Welty (University of California; San Francisco); Bharat Parekh and Dawn K. Smith (US Centers for Disease Control); Lusine Ghazaryan (USAID); Vani Vannappagari (ViiV Healthcare); Michael P. Busch (Vitalant Research Institute); Sally Hodder (West Virginia Clinical and Translational Science Institute).

Placebo-controlled trials are no longer ethical.[8] Traditional active comparator superiority designs are becoming more difficult with each generation of new products as the SOC PrEP options achieve high degrees of efficacy.[8,18,20] Noninferiority designs rely on the constancy in treatment effect over time, requiring solid data demonstrating superiority of the standard-of-care versus placebo for the population under study to allow establishing appropriate non-inferiority margins.

FUNDING

The Forum for Collaborative Research received project funding from the Bill and Melinda Gates Foundation (award number: INV-010392/OPP1199945), and is also supported by AbbVie, Gilead, Merck, Janssen, Monogram Biosciences, ViiV Healthcare, Abbott Diagnostics,

and Quest Diagnostics. Deborah Donnell, Fei Gao, and James Hughes received funding support from National Institutes of Health grant number UM1AI068635.

CONFLICT OF INTEREST

E.G. has received consulting income and research support from Sedia Biosciences Corporation and consulting income from Gilead Sciences. A.C. is an employee and stockholder, ViiV Healthcare. M.D. is an employee and stockholder, Gilead Sciences. A.D. has received research support from Gilead Sciences and Seagen. D.G. has consulted for Gilead Sciences and has served on an Advisory Board for Merck & Co., Inc. M.R. is an employee and stockholder, Merck & Co., Inc. All other authors declare no competing interests for this work.

1. Joshi, K. *et al*. Declining HIV incidence in sub-Saharan Africa: a systematic review and meta-analysis of empiric data. *J. Int. AIDS Soc.* **24**(10), e25818 (2021).
2. Rwibasira, G.N. *et al*. Recent infections among individuals with a new HIV diagnosis in Rwanda, 2018-2020. *PLoS One* **16**(11), e0259708 (2021).
3. Mosha, N.R. *et al*. The prevalence and incidence of HIV in the ART era (2006-2016) in North West Tanzania. *Int. J. STD AIDS* **33**, 346 (2022).
4. UNAIDS. Global AIDS Update (2020) [cited 2022 April 4, 2022].
5. UNAIDS 2021. April 4, 2022.
6. Baugher, A.R. *et al*. Racial, ethnic, and gender disparities in awareness of preexposure prophylaxis among HIV-negative heterosexually active adults at increased risk for HIV infection – 23 Urban Areas, United States, 2019. *MMWR Morb. Mortal. Wkly Rep.* **70**(47), 1635–1639 (2021).
7. Corey, L. *et al*. Two randomized trials of neutralizing antibodies to prevent HIV-1 acquisition. *N. Engl. J. Med.* **384**(11), 1003–1014 (2021).
8. van der Graaf, R., Reis, A. & Godfrey-Faussett, P. Revised UNAIDS/WHO ethical guidance for HIV prevention trials. *JAMA* **325**(17), 1719–1720 (2021).
9. Miller, V. The forum for collaborative HIV research: a model for an integrated and inclusive approach to clinical research and drug development. *Clin. Pharmacol. Ther.* **86**(3), 332–335 (2009).
10. Hutchison, C., Kwong, A., Ray, S., Struble, K., Swan, T. & Miller, V. Accelerating drug development through collaboration: the Hepatitis C Drug Development Advisory Group. *Clin. Pharmacol. Ther.* **96**(2), 162–165 (2014).
11. Cutrell, A. *et al*. HIV prevention trial design in an era of effective pre-exposure prophylaxis. *HIV Clin. Trials* **18**(5–6), 177–188 (2017).
12. Forum for Colalborative Research. Protocol design considerations: Analyses for efficacy. An in-depth webinar report (2021) [cited 2022 November 25, 2022].
13. ICH. Choice of Control Group and Related Issues (2001).
14. Eichler, H.G. *et al*. "Threshold-crossing": A useful way to establish the counterfactual in clinical trials? *Clin. Pharmacol. Ther.* **100**(6), 699–712 (2016).
15. Mullick, C. & Murray, J. Correlations between human immunodeficiency virus (HIV) infection and rectal gonorrhea incidence in men who have sex with men: implications for future HIV preexposure prophylaxis trials. *J. Infect. Dis.* **221**(2), 214–217 (2020).
16. Donnell, D. *et al*. Association between rectal gonorrhoea and HIV incidence in men who have sex with men: a meta-analysis. *Sex. Transm. Infect.* **98**(7), 492–496 (2022).
17. Glidden, D.V. *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. Int. AIDS Soc.* **24**(5), e25744 (2021).
18. Mayer, K.H. *et al*. 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. *Lancet* **396**(10246), 239–254 (2020).
19. Nkambule, R. *et al*. HIV incidence, viremia, and the national response in Eswatini: Two sequential population-based surveys. *PLoS One* **16**(12), e0260892 (2021).
20. Landovitz, R.J. *et al*. Cabotegravir for HIV prevention in cisgender men and transgender women. *N. Engl. J. Med.* **385**(7), 595–608 (2021).
21. Grant, R.M. *et al*. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. *N. Engl. J. Med.* **363**(27), 2587–2599 (2010).
22. Choopanya, K. *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* **381**(9883), 2083–2090 (2013).
23. Baeten, J.M. *et al*. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. *N. Engl. J. Med.* **367**(5), 399–410 (2012).
24. Antoni, G. *et al*. On-demand pre-exposure prophylaxis with tenofovir disoproxil fumarate plus emtricitabine among men who have sex with men with less frequent sexual intercourse: a post-hoc analysis of the ANRS IPERGAY trial. *Lancet HIV* **7**(2), e113–e120 (2020).
25. McCormack, S. *et al*. Pre-exposure prophylaxis to prevent the acquisition of HIV-1 infection (PROUD): effectiveness results from the pilot phase of a pragmatic open-label randomised trial. *Lancet* **387**(10013), 53–60 (2016).
26. U.S. Department of Health and Human Services Food and Drug Administration. *Demonstrating Substantial Evidence of Effectiveness for Human Drug and Biological Products Guidance for Industry* <https://www.fda.gov/media/133660/download> (2019). Cited March 24, 2021.
27. Brown, B.J. & Sugarman, J. Why ethics guidance needs to be updated for contemporary HIV prevention research. *J. Int. AIDS Soc.* **23**(5), e25500 (2020).
28. Molina, J.-M. *et al*. Efficacy, safety, and effect on sexual behaviour of on-demand pre-exposure prophylaxis for HIV in men who have sex with men: an observational cohort study. *The Lancet HIV* **4**(9), e402–e410 (2017).
29. U.S. Department of Health and Human Services Food and Drug Administration. *Guidance for Industry: E 10 Choice of Control Group and Related Issues in Clinical Trials* <https://www.fda.gov/media/71349/download> (2001). Cited March 24, 2021.
30. U.S. Department of Health and Human Services Food and Drug Administration. *Establishing Effectiveness and Safety for Hormonal Drug Products Intended to Prevent Pregnancy Guidance for Industry*. <https://www.fda.gov/media/128792/download> 2019. Cited November 16, 2021.
31. Murphy, G. *et al*. Moving towards a reliable HIV incidence test – current status, resources available, future directions and challenges ahead. *Epidemiol. Infect.* **145**(5), 925–941 (2017).
32. Parekh, B.S. & McDougal, J.S. Application of laboratory methods for estimation of HIV-1 incidence. *Indian J. Med. Res.* **121**(4), 510–518 (2005).
33. Rutherford, G.W., Schwarcz, S.K. & McFarland, W. Surveillance for incident HIV infection: new technology and new opportunities. *J. Acquir. Immune Defic. Syndr.* **25**(Suppl 2), S115–S119 (2000).
34. Smolen-Dzirba, J. & Wasik, T.J. Current and future assays for identifying recent HIV infections at the population level. *Med. Sci. Monit.* **17**(5), RA124-33 (2011).
35. Hallett, T.B. Estimating the HIV incidence rate: recent and future developments. *Curr. Opin. HIV AIDS* **6**(2), 102–107 (2011).
36. Guy, R. *et al*. Accuracy of serological assays for detection of recent infection with HIV and estimation of population

incidence: a systematic review. *Lancet Infect. Dis.* **9**(12), 747–759 (2009).
37. McDougal, J.S. *et al*. Surveillance for HIV-1 incidence using tests for recent infection in resource-constrained countries. *AIDS* **19**(Suppl 2), S25–S30 (2005).
38. Kin-On Lau, J., Murdock, N., Murray, J., Justman, J., Parkin, N. & Miller, V. A systematic review of limiting antigen avidity enzyme immunoassay for detection of recent HIV-1 infection to expand supported applications. *J. Virus Erad.* **8**(3), 100085 (2022).
39. Facente, S.N. *et al*. Use of HIV recency assays for HIV incidence estimation and other surveillance use cases: systematic review. *JMIR Public Health Surveill.* **8**(3), e34410 (2022).
40. Brookmeyer, R. & Quinn, T.C. Estimation of current human immunodeficiency virus incidence rates from a cross-sectional survey using early diagnostic tests. *Am. J. Epidemiol.* **141**(2), 166–172 (1995).
41. Kassanjee, R., McWalter, T.A., Barnighausen, T. & Welte, A. A new general biomarker-based incidence estimator. *Epidemiology* **23**(5), 721–728 (2012).
42. Kassanjee, R., McWalter, T.A. & Welte, A. Short Communication: Defining optimality of a test for recent infection for HIV incidence surveillance. *AIDS Res. Hum. Retrovir.* **30**(1), 45–49 (2014).
43. UNAIDS. *Technical update on HIV incidence assays for surveillance and monitoring purposes* <https://www.unaids.org/en/resources/documents/2015/HIVincidenceassayssurveillancemonitoring> (2015). Cited April 15, 2015.
44. Grebe, E., Facente, S.N., Bingham, J., Pilcher, C.D. *et al*. Interpreting HIV diagnostic histories into infection time estimates: analytical framework and online tool. *BMC Infect. Dis.* **19**(1), 894 (2019).
45. FIND. *Target product profile for tests for recent HIV infection* <https://www.finddx.org/wp-content/uploads/2019/03/HIV-Incidence-TPP-FIND-2017.pdf> (2017). Cited January 15, 2021.
46. Fogel, J.M. *et al*. Brief Report: Impact of Early Antiretroviral Therapy on the Performance of HIV Rapid Tests and HIV Incidence Assays. *J. Acquir. Immune Defic. Syndr.* **75**(4), 426–430 (2017).
47. Kassanjee, R. *et al*. Viral load criteria and threshold optimization to improve HIV incidence assay characteristics. *AIDS* **30**(15), 2361–2371 (2016).
48. Kassanjee, R. *et al*. Independent assessment of candidate HIV incidence assays on specimens in the CEPHIA repository. *AIDS* **28**(16), 2439–2449 (2014).
49. Kassanjee, R. *et al*. Seroconverting blood donors as a resource for characterising and optimising recent infection testing algorithms for incidence estimation. *PLoS One* **6**(6), e20027 (2011).
50. Duong, Y.T. *et al*. Recalibration of the limiting antigen avidity EIA to determine mean duration of recent infection in divergent HIV-1 subtypes. *PLoS One* **10**(2), e0114947 (2015).
51. Duong, Y.T., Qiu, M., De, A.K., Jackson, K. *et al*. Detection of recent HIV-1 infection using a new limiting-antigen avidity assay: potential for HIV-1 incidence estimates and avidity maturation studies. *PLoS One* **7**(3), e33328 (2012).
52. Parekh, B.S. *et al*. Determination of mean recency period for estimation of HIV type 1 Incidence with the BED-capture EIA in persons infected with diverse subtypes. *AIDS Res. Hum. Retrovir.* **27**(3), 265–273 (2011).
53. Seaton, K.E., Vandergrift, N.A., Deal, A.W., Rountree, W. *et al*. Computational analysis of antibody dynamics identifies recent HIV-1 infection. *JCI Insight* **2**(24), e94355 (2017).
54. Laeyendecker, O., Kulich, M., Donnell, D., Komarek, A. *et al*. Development of methods for cross-sectional HIV incidence estimation in a large, community randomized trial. *PLoS One* **8**(11), e78818 (2013).
55. Laeyendecker, O. *et al*. Specificity of four laboratory approaches for cross-sectional HIV incidence determination: analysis of samples from adults with known nonrecent HIV infection from five African countries. *AIDS Res. Hum. Retrovir.* **28**(10), 1177–1183 (2012).
56. Duong, Y.T. *et al*. Field validation of limiting-antigen avidity enzyme immunoassay to estimate HIV-1 incidence in cross-sectional survey in Swaziland. *AIDS Res. Hum. Retrovir.* **35**(10), 896–905 (2019).
57. Low, A. *et al*. Correlates of HIV infection in adolescent girls and young women in Lesotho: results from a population-based survey. *Lancet HIV* **6**(9), e613–e622 (2019).
58. Gonese, E. *et al*. Comparison of HIV incidence in the Zimbabwe population-based HIV impact assessment survey (2015-2016) with modeled estimates: progress toward epidemic control. *AIDS Res. Hum. Retrovir.* **36**(8), 656–662 (2020).
59. Justman, J.E., Mugurungi, O. & El-Sadr, W.M. HIV population surveys – bringing precision to the global response. *N. Engl. J. Med.* **378**(20), 1859–1861 (2018).
60. Rice, B.D., de Wit, M., Welty, S., Risher, K. *et al*. Can HIV recent infection surveillance help us better understand where primary prevention efforts should be targeted? Results of three pilots integrating a recent infection testing algorithm into routine programme activities in Kenya and Zimbabwe. *J. Int. AIDS Soc.* **23**(Suppl 3), e25513 (2020).
61. Farahani, M., Radin, E., Saito, S., Sachathep, K. *et al*. Population viral load, viremia and recent HIV-1 infections: findings from population-based HIV impact assessments (PHIAs) in Zimbabwe, Malawi, and Zambia. *J. Acquir. Immune Defic. Syndr.* **87**, S81–S88 (2021).
62. Klock, E. *et al*. Validation of population-level HIV-1 incidence estimation by cross-sectional incidence assays in the HPTN 071 (PopART) trial. *J. Int. AIDS Soc.* **24**(12), e25830 (2021).
63. Keating, S.M. *et al*. Development of an international external quality assurance program for HIV-1 incidence using the Limiting Antigen Avidity assay. *PLoS One* **14**(9), e0222290 (2019).
64. Curtis, K.A. *et al*. Evaluation of the Abbott ARCHITECT HIV Ag/Ab combo assay for determining recent HIV-1 infection. *PLoS One* **16**(7), e0242641 (2021).
65. Hassan, J., Moran, J., Murphy, G., Mason, O., Connell, J. & de Gascun, C. Discrimination between recent and non-recent HIV infections using routine diagnostic serological assays. *Med. Microbiol. Immunol.* **208**(5), 693–702 (2019).
66. Grebe, E. *et al*. Infection staging and incidence surveillance applications of high dynamic range diagnostic immuno-assay platforms. *J. Acquir. Immune Defic. Syndr.* **76**(5), 547–555 (2017).
67. Sempa, J.B. *et al*. Performance comparison of the Maxim and Sedia Limiting Antigen Avidity assays for HIV incidence surveillance. *PLoS One* **14**(7), e0220345 (2019).
68. Hemelaar, J., Gouws, E., Ghys, P.D. & Osmanov, S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* **20**(16), W13–W23 (2006).
69. Lynch, B.A. *et al*. Short communication: false recent ratio of the limiting-antigen avidity assay and viral load testing algorithm among cameroonians with long-term HIV infection. *AIDS Res. Hum. Retrovir.* **33**(11), 1114–1116 (2017).
70. Mullis, C.E. *et al*. Differential specificity of HIV incidence assays in HIV subtypes A and D-infected individuals from Rakai, Uganda. *AIDS Res Hum Retroviruses* **29**(8), 1146–1150 (2013).
71. Longosz, A.F. *et al*. Comparison of antibody responses to HIV infection in Ugandan women infected with HIV subtypes A and D. *AIDS Res. Hum. Retrovir.* **31**(4), 421–427 (2015).
72. Longosz, A.F. *et al*. Immune responses in Ugandan women infected with subtypes A and D HIV using the BED capture immunoassay and an antibody avidity assay. *J. Acquir. Immune Defic. Syndr.* **65**(4), 390–396 (2014).
73. Longosz, A.F. *et al*. Impact of HIV subtype on performance of the limiting antigen-avidity enzyme immunoassay, the bio-rad avidity assay, and the BED capture immunoassay in Rakai, Uganda. *AIDS Res Hum Retroviruses* **30**(4), 339–344 (2014).
74. de Wit, M.M. *et al*. Experiences and lessons learned from the real-world implementation of an HIV recent infection testing algorithm in three routine service-delivery settings in Kenya and Zimbabwe. *BMC Health Serv. Res.* **21**(1), 596 (2021).
75. FIND. *Required sample size for power* (*baseline survey and cohort*). <https://finddx.shinyapps.io/sample_size_baseline_and_cohort/> (2018). Cited February 18, 2021.
76. Gao, F., Glidden, D.V., Hughes, J.P. & Donnell, D. *Sample Size Calculation for Active-Arm Trial with Counterfactual Incidence Based*

on Recency Assay. arXiv <https://arxiv.org/abs/2011.00725> (2020). Cited February 18, 2021.
77. Abdool Karim, Q., Kharsany, A.B., Frohlich, J.A., Werner, L. *et al*. HIV incidence in young girls in KwaZulu-Natal, South Africa – public health imperative for their inclusion in HIV biomedical intervention trials. *AIDS Behav.* **16**(7), 1870–1876 (2012).
78. Lane, T., Osmand, T., Marr, A., Struthers, H., McIntyre, J.A. & Shade, S.B. Brief report: High HIV Incidence in a South African Community of Men Who Have Sex With Men: Results From the Mpumalanga Men's Study, 2012-2015. *J. Acquir. Immune Defic. Syndr.* **73**(5), 609–611 (2016).
79. Evidence for Contraceptive Options and HIV Outcomes (ECHO) Trial Consortium. HIV incidence among women using intramuscular depot medroxyprogesterone acetate, a copper intrauterine device, or a levonorgestrel implant for contraception: a randomised, multicentre, open-label trial. *Lancet* **394**(10195), 303–313 (2019).
80. Van Damme, L., Corneli, A., Ahmed, K., Agot, K. *et al*. Preexposure prophylaxis for HIV infection among African women. *N. Engl. J. Med.* **367**(5), 411–422 (2012).
81. Thigpen, M.C. *et al*. Antiretroviral preexposure prophylaxis for heterosexual HIV transmission in Botswana. *N. Engl. J. Med.* **367**(5), 423–434 (2012).
82. ClinicalTrials.gov. *Evaluating the Safety and Efficacy of Long-Acting Injectable Cabotegravir Compared to Daily Oral TDF/FTC for Pre-Exposure Prophylaxis in HIV-Uninfected Women (HPTN 084)* <https://clinicaltrials.gov/ct2/show/NCT03164564> (2017). Cited March 4, 2021.
83. ClinicalTrials.gov. *Safety and Efficacy Study of Injectable Cabotegravir Compared to Daily Oral Tenofovir Disoproxil Fumarate/Emtricitabine (TDF/FTC), For Pre-Exposure Prophylaxis in HIV-Uninfected Cisgender Men and Transgender Women Who Have Sex With Men (HPTN 083)* <https://clinicaltrials.gov/ct2/show/NCT02720094> (2016). Cited March 4, 2021.
84. Republic of South Africa Department of Health. *National HIV Testing Services: Policy* <https://sahivsoc.org/Files/HTS%20Policy%2028%20July%20final%20copy.pdf> (2016). Cited April 9, 2021.
85. Laeyendecker, O. *et al*. Antibody maturation in women who acquire HIV infection while using antiretroviral preexposure prophylaxis. *J. Infect. Dis.* **212**(5), 754–759 (2015).
86. Curtis, K.A. *et al*. Delayed maturation of antibody avidity but not seroconversion in rhesus macaques infected with simian HIV during oral pre-exposure prophylaxis. *J. Acquir. Immune Defic. Syndr.* **57**(5), 355–362 (2011).
87. Donnell, D. *et al*. HIV protective efficacy and correlates of tenofovir blood concentrations in a clinical trial of PrEP for HIV prevention. *J. Acquir. Immune Defic. Syndr.* **66**(3), 340–348 (2014).