External Comparison of Effectiveness of Ibalizumab in Clinical trials vs. Other HTE Regimens in OPERA

Protocol

August 4, 2022

Authors:

Laurence Brunet, Epividian

Jennifer Fusco, Epividian

Gregory Fusco, Epividian

Brandon Cash, TheraTechnologies

Lydia Walder, FIECON

Andrew Laws, FIECON

Table of Contents

List of Tables

List of Figures

List of Equations

Executive Summary

Background & Rationale

HIV viral resistance to antiretrovirals (ARV), whether transmitted or acquired following long-term antiretroviral therapy (ART) exposure, can greatly limit therapeutic options.[1-8] As a result, a subset of people with HIV (PWH) referred to as heavily treatment-experienced (HTE) require highly tailored ART regimens with less common ARV combinations.[9] In the US, HTE prevalence has been estimated to range from 2% to 14% in 2016-2017.[10, 11] The prognosis for HTE PWH may be less favorable compared to less experienced PWH. Notably, multidrug resistance (MDR) has been associated with a higher risk of death, AIDS-defining events (ADE) and viral rebounds.[9, 12-14] Though the development of new ARVs has increased the number of treatment options available, there are still individuals who cannot achieve complete virologic suppression and remain at risk for disease progression.[15]

The US Department of Health and Human Services (DHHS) HIV treatment guideline suggest that IBA may be considered in the context of MDR without fully active ART options, namely among PWH with ongoing detectable viremia and lacking sufficient options to build a fully suppressive regimen.[16] However, clinical trials assessing IBA efficacy did not include an active comparison arm in which participants did not receive IBA. An external comparison analysis has been conducted between IBA + optimized background regimen (OBR) from clinical trials (TMB-202, TMB-301/311), and non-IBA OBR from a European registry. The OPERA study will improve upon this European registry study with a larger sample size and the use of controls from the same geographic region as the trial participants.

It is important to assess the comparative effectiveness of IBA compared to other regimens used in the routine clinical care of HTE PWH with MDR HIV. The use of external controls can provide valuable information and context to interpret the results of clinical trials when randomization to a control arm cannot be performed.[17] Using comparable populations and larger sample sizes may reduce bias and improve the power of analyses. The OPERA (Observational Pharmaco-Epidemiology Research & Analysis) cohort, a large US electronic health record (EHR) database, is well suited for this as the OPERA and trials populations arose from the same geographic location (i.e., US), and the number of OPERA PWH potentially meeting inclusion criteria is high. Therefore, an external comparison of IBA+OBR from trials to non-IBA OBR in routine clinical care in the OPERA cohort may confirm results from prior studies suggesting viral control benefits of IBA+OBR.

Study Objectives

Primary Objective:

To assess the virologic effectiveness of 800 mg IBA every 2 weeks + OBR (TMB-202 & TMB-301/311 trials), compared to non-IBA OBR in routine clinical care (OPERA cohort), among HTE PWH.

Sensitivity Objective 1:

To assess the robustness of results from the primary objective to the viremia threshold for inclusion in the OPERA non-IBA OBR group (main: VL >1,000 copies/mL vs. sensitivity: VL >200 copies/mL).

Sensitivity Objective 2:

To replicate previous analyses of the European registry as closely as possible, taking into consideration the differences between the registry and OPERA datasets

Data sources

Data from the phase 2 clinical trial TMB-202, as well as the phase 3 trial TMB-301 and its extension TMB-311 will be used for the treatment arm of the study (IBA+OBR). The OPERA cohort, which includes complete EHR data from over 143,000 PWH across 142 separate locations throughout the US, will be used for the control arm (non-IBA OBR).

Study Design

All TMB-202 and TMB-301/311 trial participants randomized to 800 mg IBA every 2 weeks with or without a loading dose will be included in the main analyses. From the OPERA cohort, adult PWH with resistance from three ARV classes will be considered for inclusion (see section 3.2.1 for more details). OPERA PWH initiating a new non-IBA OBR between 01JAN2008 and 31DEC2020 will be included; person-time will be censored at the first of (a) initiation of a select potent ARV, (b) >45 days without ART, (c) 12 months after last clinical contact, (d) death, or (e) study end (i.e., extract date of the dataset; TBD). Viral suppression (i.e., <200 copies/mL) and undetectable viral load (i.e., <50 copies/ml) will be assessed at each of the timepoints of interest (i.e., weeks 24, 48, 60 and 96), using the closest measurement to the timepoint within a ±6-week window.

Statistical Analyses

Multiple imputation will be used to ensure that eligible OPERA participants can contribute to all timepoint analyses even if their routine clinical care resulted in viral loads being measured outside of the ranges defined in the protocol. To prevent the introduction of a bias towards the null, imputation will be performed in the OPERA non-IBA OBR group separately.[18] 95% confidence intervals will be obtained through bootstrapping.

First, the data will be sampled with replacement to create 500 bootstrap samples. Within each bootstrap sample, multiple imputation will be implemented using chained equations, producing 10 imputed datasets per bootstrap sample.

Within each imputed dataset, analyses will be performed to obtain an estimate at each timepoint for all 5,000 datasets (i.e., 500 bootstrap samples * 10 imputed datasets). Risk ratios for both viral suppression and undetectability will be estimated as a measure of comparative virologic effectiveness of IBA+OBR vs. non-IBA OBR at each timepoint. The *average treatment effect in the treated* (ATT) will be estimated using standardized mortality ratio (SMR)-weighting, a form of propensity score (PS)-weighting. SMR-weighting ensures that the distribution of baseline characteristics is the same in the control population as in the treated population.[20] In other words, the external control population will be weighted to look like the treated population. Logistic regression will be used to obtain the PS and will include baseline values for important predictors of regimen selection. SMR weights are set to 1 in the treated population (i.e., TMB trials IBA+OBR) and to PS/(1-PS) in the control population (i.e., OPERA non-IBA OBR).[20] The final set of variables included in the PS model will be determined by the distribution of weights, as well as the balance of covariates across groups before and after applying SMR weights.

Estimates of risk will be derived from SMR-weighted Kaplan-Meier curves. The cumulative probability of viral suppression or undetectability will be estimated at each timepoint of interest for each arm, from which risk ratios will be calculated. Estimates will be averaged across all datasets to obtain the combined estimate and 95% confidence intervals will be derived from the distribution of all estimates, using the 2.5^{th} & 97.5^{th} percentiles.

These analyses will be repeated for sensitivity objective 1 after restricting the OPERA non-IBA OBR group to PWH with a VL >1,000 copies/mL. For sensitivity objective 2, analyses performed using the European registry data will be replicated as closely as possible, taking into consideration the differences in design between the registry and OPERA datasets.

1. Background and Rationale

In 2020, over 37 million people were living with HIV infection worldwide, and 680,000 people died from AIDS.[21] In regions of the world like the US with access to antiretroviral therapy (ART), the development of antiretrovirals (ARV) has prolonged life expectancy for people with HIV (PWH).[22] However, long-term exposure to ART can eventually lead to fewer therapeutic options in multiple ARV classes.[1, 2] This can be attributed to multiple factors including poor tolerability, drug toxicity, avoidance of drug-drug interactions, and viral resistance.[2-6] As a result, a subset of PWH require highly tailored ART regimens with less common ARV combinations; these individuals are often referred to as heavily treatment-experienced (HTE).[9] Moreover, while transmitted resistance has decreased over calendar time, it remains a concern among ART-naïve PWH; some ART-naïve individuals may thus also have limited treatment options.[7, 8]

HTE prevalence varies depending on the setting and the definition used. In the US, HTE prevalence has been estimated to range from 2% to 14% in 2016-2017.[10, 11] In Europe, HTE prevalence increased from 6% in 2010 to 9% in 2016; important geographic variation has also been observed, from 1% in Eastern Europe to 16% in Western/Central Europe.[23]

The prognosis for HTE PWH appears to be less favorable compared to less experienced PWH. Increased risk of death and AIDS-defining events (ADE) have been reported among PWH with multidrug resistance (MDR).[9, 12, 13] Moreover, triple ARV class failure has been associated with a higher rate of viral rebound (i.e., two consecutive viral loads > 400 copies/mL or one viral load > 400 copies/mL followed by initiation of at least 2 new ARVs).[14] Published evidence suggests that mortality in people with HIV is associated with incomplete viral suppression. Studies investigating the effect of detectable HIV RNA after antiretroviral therapy and mortality rates in people with HIV identified that incomplete suppression of HIV RNA after receiving ART is associated with significant 10-year all-cause mortality risk.[24, 25]

Complex ART regimens, which are characteristic of treatments for HTE PWH, may both lack sufficient efficacy and have adverse safety and tolerability profiles, making disease and toxicity management even more difficult.[15, 26, 27] Though the development of new ARVs has increased the number of treatment options available for HTE PWH, there are still individuals who cannot achieve complete virologic suppression and remain at risk for disease progression.[15] The US Department of Health and Human Services (DHHS) HIV treatment guideline suggest that ibalizumab (IBA) may be considered in the context of MDR without fully active ART options, namely among PWH with ongoing detectable viremia and lacking sufficient options to build a fully suppressive regimen.[16]

The TMB-202 trial was a phase 2b study conducted between 2008-2010 in the US and Taiwan, in which 113 patients were randomized to receive either 800mg IBA every two weeks (n=59) or 2000mg IBA every four weeks (n=54), in addition to an optimized background regimen (OBR).[28] The TMB-301 trial and its extension TMB-311 were phase 3 single arm studies of 40 patients (34 in the US, 4 in Taiwan, 2 in Puerto Rico). All received a loading dose of 2000mg IBA with their failing regimen, followed by 800mg IBA every two

weeks with an OBR.[29] These trials did not include a comparison arm in which participants did not receive IBA. Indeed, given that the trials were restricted to failing individuals (i.e., viral load >1,000 copies/mL) with limited susceptibility to other approved ARVs, withholding IBA from some trial participants was deemed unethical. It is therefore important to assess the comparative effectiveness of IBA compared to other regimens used in routine clinical care of people with MDR HIV.

Such an external comparison analysis has been conducted against a European Registry, an observational, prospective, multi-center study. The OPERA study will improve upon this European registry study with a larger sample size and the use of controls from the same geographic region as the trial participants.

The use of external controls can provide valuable information and context to interpret the results of clinical trials when randomization to a control arm cannot be performed.[17] However, careful consideration must be given to biases that may be introduced in such analyses. Regional bias, in which patient outcomes differ based on the geographic region due to differences in standard of care or demographic characteristics, can be reduced by selecting external controls in the same region where the trial was conducted.[17] The Observational Pharmaco-Epidemiology Research & Analysis (OPERA) cohort is a large observational database electronic health record (EHR) from over 143,000 PWH across the US.

It is important to assess the comparative effectiveness of IBA compared to other regimens used in the routine clinical care of HTE PWH with MDR HIV. External comparisons of trial participants and PWH in routine clinical care may help assess the comparative effectiveness of IBA+OBR. Using a comparable external control and larger sample sizes may reduce bias and improve the power of such indirect comparisons. Therefore, an external comparison of IBA+OBR from the TMB-202 and TMB-301/311 trials and non-IBA OBR in routine clinical care in the OPERA cohort may confirm results from prior studies suggesting viral control benefits of IBA+OBR.

2.Study Objectives

Primary Objective

I. To assess the virologic effectiveness of 800mg IBA every 2 weeks + OBR (TMB-202 & TMB-301/311 trials), compared to other ART regimens in routine clinical care (OPERA cohort), among HTE PWH.

- External comparison (IBA+OBR in clinical trials vs. non-IBA OBR in OPERA)
- Timepoint analysis
- Multiple imputation of missing viral loads in OPERA
- PS standardized mortality ratio-weighting

Sensitivity Objectives

I. To assess the robustness of results from the primary objective to the viremia threshold for inclusion in the OPERA non-IBA OBR group (main: VL >1,000 copies/mL vs. sensitivity: VL >200 copies/mL).

II. To replicate previous analyses of the European registry as closely as possible, taking into consideration the differences in design between the registry and OPERA datasets

- External comparison (IBA+OBR in clinical trials vs. non-IBA OBR in OPERA)
- Timepoint analysis
- Complete case analysis
- Propensity score matching (PS including baseline VL, CD4 cell count, number of potent ARVs scaled by the level of resistance, and the calendar period score)

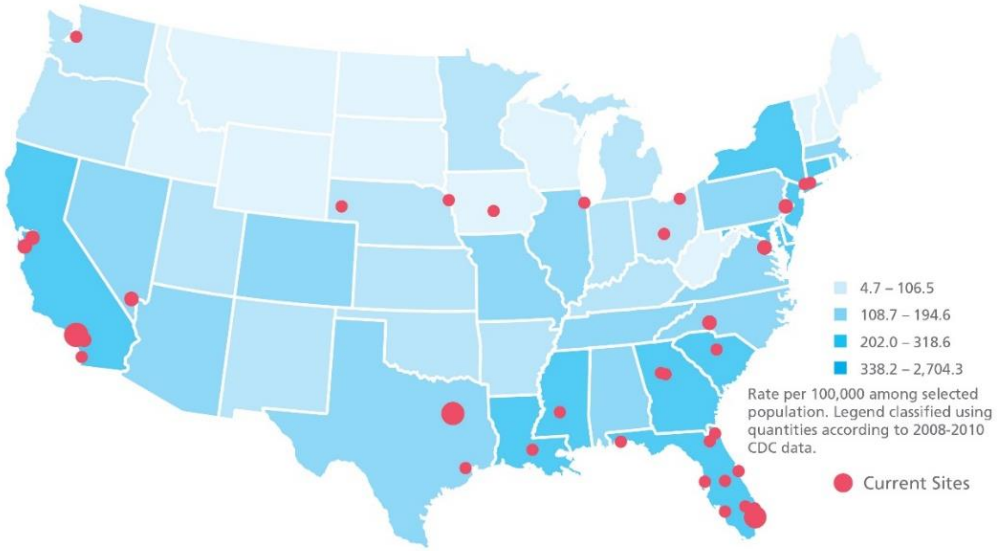
3. Study Design

3.1. Data sources

3.1.1. OPERA Cohort

The OPERA® (Observational Pharmaco-Epidemiology Research & Analysis) database and research network is a multi-site observational database built from the complete patient health records managed in EHR systems from more than 400 participating healthcare providers at 142 separate locations throughout the US (see coverage map below, Figure 1). Through their membership in OPERA, medical practices meet the Centers for Medicare & Medicaid Services (CMS) MIPS Incentive Program for Integration with a Specialized Registry. OPERA-participating physicians and ancillary healthcare providers have documented the care of over 900,000 patients in their EHRs, including over 143,000 PWH (~20% women), representing 14% of all PWH linked to care in the US.[31] The OPERA database is refreshed from these EHR systems at each clinic daily, providing up-to-date data to both clinicians and researchers. In total, there are more than 12.7 million documented prospective visits in the EHR systems for PWH and over 4.3 million prescriptions written for ART medications. The average duration of follow-up (i.e., years of documenting patient visits prospectively in the EHR) is 4.7 years and there are over 20,000 PWH who have 10 years or more of follow-up.

Figure 1. US map of OPERA HIV+ population & CDC (2010) state-by-state estimates



3.1.2. Clinical Trials

Data from the phase 2 clinical trial TMB-202, as well as from the phase 3 trials TMB-301 and TMB-311 will be used.

TMB-202 trial

The TMB-202 trial (NCT00784147)[28] was a phase 2b, randomized, double-blinded, multi-center, dose-response study of IBA+OBR in ART-experienced patients infected with HIV-1. The 48-week study was amended to 24-weeks. A total of 113 patients were enrolled between 2008 and 2010 in the US and Taiwan. Patients were randomized between IBA 800mg IV every two weeks (n=59) and IBA 2000mg IV every four weeks (n=54).

Trial Inclusion Criteria (criteria applicable to OPERA are underlined)

1. Are capable of understanding and have voluntarily signed the informed consent document
2. Have documented HIV-1 infection
3. Have no acquired immunodeficiency syndrome (AIDS)-defining events in the 3 months before Screening, other than cutaneous Kaposi's sarcoma or wasting syndrome due to HIV
4. Are able and willing to comply with all protocol requirements and procedures
5. Are 18 years of age or older
6. Have a life expectancy that is >6 months.
7. Have a viral load >1,000 copies/mL and documented decreased susceptibility to at least one NRTI, one NNRTI, and one PI, as measured by resistance testing
8. Are receiving a stable HAART for at least 8 weeks before Screening and are willing to continue that regimen until the Baseline Visit, OR (in the past 8 weeks) have failed and are off therapy and are willing to stay off therapy until the Baseline Visit
9. Have viral sensitivity/susceptibility to at least one agent as determined by the Screening resistance test (OSS criteria) and be willing and able to be treated with at least one agent to which the patient's viral isolate is sensitive/susceptible according to the screening resistance test
10. If sexually active, are willing to use an effective method of contraception during the study and for 30 days after the last administration of the study drug

Trial Exclusion Criteria (criteria applicable to OPERA are underlined)

1. Any active AIDS-defining illness per Category C conditions according to the CDC Classification System for HIV Infection13 (see Appendix D), with the following exceptions: cutaneous Kaposi's sarcoma and wasting syndrome due to HIV
2. Any significant diseases (other than HIV-1 infection) or clinically significant findings, including psychiatric and behavioral problems, determined from screening, medical history and/or physical examination that, in the investigator's opinion, would preclude the patient from participating in this study
3. Any acute illness within 1 week before the initial administration of study drug
4. Any active infection secondary to HIV requiring acute therapy; however, patients that require maintenance therapy (i.e., secondary prophylaxis for opportunistic infections) will be eligible for the study.
5. Any immunomodulating therapy (including interferon), systemic steroids, or systemic chemotherapy within 12 weeks before randomization
6. Any investigational therapy within 30 days before randomization, except for those HIV-agents available in expanded-access programs
7. Any prior exposure to ibalizumab (formerly TNX-355 and Hu5A8)
8. Any vaccination within 21 days before randomization
9. Any female patient who either is pregnant, intends to become pregnant, or is currently breastfeeding

TMB-301/TMB-311 trials

The TMB-301/311 trials (NCT02475629, NCT02707861)[29, 32] were phase 3, single arm, multi-center studies of IBA+OBR in ART-experienced patients infected with MDR HIV-1. TMB-301 was a 24-week study and TMB-311 extended its follow-up to up to 96 weeks. All participants received 2000mg IBA IV (loading dose) as they maintained their current regimen, followed 14 days later by 800mg IBA IV every two weeks + OBR. A total of 40 patients were enrolled in TMB-301 (34 in the US, 4 in Taiwan, 2 in Puerto Rico), of which 27 US patients were enrolled into the extension study TMB-311 and continued to receive IBA+OBR at the dosage assigned in TMB-301. There were four deaths during the 24-week period of TMB-301, and two deaths during the TMB-311 period.

Inclusion Criteria (criteria applicable to OPERA are underlined)

1. Are capable of understanding and have voluntarily signed the informed consent document
2. Have documented HIV-1 infection
3. Have no AIDS-defining events in the 3 months before Screening, other than cutaneous Kaposi's sarcoma or wasting syndrome due to HIV
4. Are able and willing to comply with all protocol requirements and procedures
5. Are 18 years of age or older
6. Have a life expectancy that is >6 months
7. Have a viral load >1,000 copies/mL and documented resistance to at least one antiretroviral medication from each of three classes of antiretroviral medications as measured by resistance testing
8. Have a history of at least 6 months on antiretroviral treatment
9. Are receiving a stable highly active antiretroviral regimen for at least 8 weeks before screening and are willing to continue that regimen until Day 14, OR (in the past 8 weeks) have failed and are off therapy and are willing to stay off therapy until Day 14
10. Have full viral sensitivity/susceptibility to at least one antiretroviral agent, other than ibalizumab, as determined by the screening resistance tests and be willing and able to be treated with at least one agent to which the patient's viral isolate is fully sensitive/susceptible according to the screening resistance tests as a component of OBR[1]
11. If sexually active, are willing to use an effective method of contraception during the study and for 30 days after the last administration of the study drug

Exclusion criteria (criteria applicable to OPERA are underlined)

1. Any active AIDS-defining illness per Category C conditions according to the CDC Classification System for HIV Infection (see Appendix A), with the following exceptions: cutaneous Kaposi's sarcoma and wasting syndrome due to HIV
2. Any significant diseases (other than HIV-1 infection) or clinically significant findings, including psychiatric and behavioral problems, determined from screening, medical history and/or physical examination that, in the investigator's opinion, would preclude the patient from participating in this study
3. Any significant acute illness within 1 week before the initial administration of study drug;
4. Any active infection secondary to HIV requiring acute therapy; however, patients that require maintenance therapy (i.e., secondary prophylaxis for opportunistic infections) will be eligible for the study
5. Any immunomodulating therapy (including interferon), systemic steroids, or systemic chemotherapy within 12 weeks before Enrollment
6. Any prior exposure to ibalizumab (formerly TNX-355 and Hu5A8)
7. Any vaccination within 7 days prior to Day 0/Enrollment
8. Any female patient who either is pregnant, intends to become pregnant, or is currently breastfeeding

[1] 5 patients with no active ARV other than IBA were enrolled.

9. Any current alcohol or illicit drug use that, in the investigator's opinion, will interfere with the patient's ability to comply with the study schedule and protocol evaluations
10. Any previous clinically significant allergy or hypersensitivity to any excipient in the ibalizumab formulation
11. Any radiation therapy during the 28 days before the first administration of investigational medication
12. Any Grade 3 or 4 laboratory abnormality according to the Division of AIDS (DAIDS) grading scale (Appendix B), except for the following asymptomatic Grade 3 events (triglyceride elevation; or total cholesterol elevation)

3.2. Study Population

3.2.1. Inclusion Criteria

Trial IBA+OBR group

1. Participant in TMB-202 or TMB-301/311[2]
2. Received 800 mg IBA every 2 weeks, with or without a loading dose

OPERA non-IBA OBR group

1. HIV-1 infection
2. 18 years or older
3. HTE (i.e., ever on a regimen containing either DTG BID or DRV BID)[3]
4. Documented resistance to ≥1 ARV from each of three ARV classes
5. Genotype information available for all the relevant gene regions (see Error! Reference source not found.)
6. Viral load > 200 copies/mL at index regimen initiation
7. Not pregnant at index
8. No new ADE within 3 months before/on index (except cutaneous Kaposi's sarcoma or HIVAW)
9. No new cancer diagnosis within 12 weeks before/on index

[2] Participant on a FTR-containing OBR may be excluded if the OPERA non-IBA regimens do not include any FTR-containing regimens
[3] HTE definition may be expanded if a larger sample size is needed

10. No IBA prescribed prior to/with the index regimen
11. Index regimen does not include
    - Cabotegravir (CAB)[4]
    - Fostemsavir (FTR)
    - Investigational drugs
12. Viral sensitivity/susceptibility to at least one ARV in the index regimen
13. Baseline VL available (within 6 months before/at index)
14. Baseline CD4 available (within 6 months before/at index)
15. ≥1 VL measurement at any time after index

3.2.2. Study Timelines

Study timelines are mapped out for the two trials and for OPERA in Table 1.

TMB-202

- Recruitment: 2008-2010
- Follow-up: through week 24
- Index: date of first IBA infusion

TMB-301/311

- Recruitment: 2015-2018
- Follow-up: through week 96
- Index: date of first IBA infusion

[4] CAB had not yet been approved at the time and is not indicating for ART-experienced PWH who are failing their current regimen

OPERA

- Identification of study participants: 01JAN2008 to 31DEC2020
- Censoring criteria:
    a) Initiation of any of the following ARVs:
        - INSTI
        - PI
        - Fusion inhibitor (FI)
        - Gp120 inhibitor
        - IBA
        - Doravirine
        - Blinded ARV
    b) >45 days without ART
    c) 12 months after last clinical contact (i.e., telephone contact, visit, lab test, or consultation)
    d) Death
    e) Study end (i.e., extract date; TBD)
- Index: date of regimen initiation

Table 1. Alignment of study timelines for the clinical trials and OPERA

<table>
<tr><th></th><th>Week -1</th><th>Index</th><th>Week 1</th><th>Week 24</th><th>Week 48</th><th>Week 60</th><th>Week 96</th></tr>
<tr><td rowspan="2">TMB-202</td><td rowspan="2">Failing Regimen</td><td colspan="3">IBA 2000mg every 4 weeks + OBR[a]</td><td rowspan="2"></td><td rowspan="2"></td><td rowspan="2"></td></tr>
<tr><td colspan="3">IBA 800mg every 2 weeks + OBR[a]</td></tr>
<tr><td>TMB-301/311</td><td>Failing Regimen</td><td>IBA 2000mg + Failing Regimen</td><td colspan="5">IBA 800mg every 2 weeks + OBR[a]</td></tr>
<tr><td>OPERA</td><td>Failing Regimen</td><td colspan="5">Index regimen[b]</td></tr>
</table>

[a] Changes in OBR were allowed during follow-up
[b] Some changes in ARVs will be allowed to be consistent with trial designs; person-time will be censored at initiation of an INSTI, PI, fusion inhibitor, Gp120 inhibitor, IBA, doravirine or a blinded ARV

3.3. Measurements

3.3.1. Exposure Definition

IBA+OBR

Defined as either 800mg IBA every 2 weeks + OBR (TMB-202), or 800mg IBA every 2 weeks + OBR following a 2000mg IBA loading dose (TMB-301/311) (see Table 1)

vs.

Non-IBA OBR

Defined as the first HTE-indicative regimen initiated within the study period in the OPERA cohort, after resistance has been established. Recycling of ARVs from prior regimens is allowed. Viral sensitivity/susceptibility to at least one ARV in the regimen.

3.3.2. Outcome Definition

Virologic outcomes:

- Viral suppression (<200 copies/mL)
- Undetectable VL (<50 copies/mL)

Timepoints of interest:

Weeks 24, 48, 60, 96 after index (closest measurement within a ±6-week window)

3.3.3. Covariates

Covariates for the propensity score model

The following characteristics are available in the TMB-202, TMB-301/311 and OPERA datasets. These baseline covariates are candidates for inclusion in the PS model.

- Age
- Sex
- Viral load
- CD4 cell count

- Index date
- OBR
- Resistance to ARVs
- Number of potent ARV received as part of index regimen, scaled by the level of resistance a patient has to these ARVs
    - Provides a measure of the number of potent ARVs which formulates each patient's OBR, where the influence of each ARV is scaled by relative resistance to that drug class
    - Based on the number of ARVs a patient is currently receiving for which they display mutations corresponding to resistance level (RL) 1-5 [see Error! Reference source not found.]
    - For each regimen, weighting of the level of resistance will be calculated using Equation 1

Equation 1

$$Weight = \sum_{1}^{n} \frac{ARV_n \times (5 - RL)}{5} + \frac{DRV \times (4 - RL)}{4}$$

Where:

ARV_n = DTG, TDF, TAF, FTC, 3TC or ETR; 1 if present, 0 if absent
DRV = 1 if DRV present, 0 if DRV absent
RL = Resistance Level [see Error! Reference source not found.]

Table 2. Categories included in drug class-specific mutation variable

Specific Drug	B Presence of Mutations or Patterns that Confer Extremely High-Level Resistance	C Presence of Mutations or Patterns that Confer High-Level Resistance	D Presence of Mutations or Patterns that Confer Intermediate-Level Resistance	E Presence of Predictive Mutations but Activity Not Impacted	F
	Category 4 Required Mutations	Category 3 Required Mutations or Score	Category 2 Required Mutations or Score	Category 1 Required Mutations or Score	Notes
DTG	Q148H, Q148K, Q148R + any three of the following L74I, L74M T97A E138A, E138K, E138T G140A, G140C, G140S	Q148H, Q148K, Q148R + any two of the following L74I, L74M T97A E138A, E138K, E138T G140A, G140C, G140S	Q148H, Q148K, Q148R + any one of the following L74I, L74M T97A E138A, E138K, E138T G140A, G140C, G140S	Two or more of the following: T66A, T66I, T66K E92Q T97A Y143R, Y143C, Y143H V151I N155H	NA
DRV	NA	Darunavir Score ≥12 (consistent with high-level resistance)	Darunavir Score 7-11 (consistent with intermediate-level resistance)	Darunavir Score ≤6	Add the score for each of these mutations to create a total DRV score: 4 each: I50V, L76V, I84V, I54MV 3 each: V32I, L33F, I47AV, I54L, V82F 2 each: V11I, G73S, L89V
TDF, TAF	NA	Includes one of: K65R, K65N Without M184V, M184I	Includes one of: K65R, K65N With M184V or M184I	NA	M184V/I confers resistance to other drugs but increases sensitivity to TDF, TAF
FTC, 3TC	NA	Includes one of: M184V, M184I	Includes one of: K65R, K65N	NA	NA
ETR	NA	Includes one of: K101P, Y181I, Y181V OR Two of: L100I, K101EH, Y181C, G1190E, G190Q, H221Y	Includes one of: L100I, K101EH, Y181C, G190E, G190Q, H221Y	NA	NA

Covariates for the multiple imputation model

The following characteristics are available in OPERA but may or may not be available in the TMB-202 and TMB 301/311 datasets. These baseline and time-updated characteristics may be used for the multiple imputation model, as predictors of the value of viral loads at timepoints of interest.

- Age at regimen initiation
- Sex at regimen initiation
- Race
- HBV co-infection at regimen initiation
- HCV co-infection at regimen initiation
- AIDS (ever)
- Viral load at each timepoint
- Viral loads in between timepoints
- Timing of viral loads
- CD4 cell count at each timepoint
- CD4 cell count in between timepoints
- Timing of CD4 cell counts
- Index date
- ARVs in index regimen
- Resistance to ARVs
- Number of anchor agent classes exposure prior to index

4. Data Analyses

4.1. Bootstrapping

Bootstrapping will be employed to obtain confidence intervals around the study estimates. First, the data will be sampled with replacement to create 500 bootstrap samples. Within each bootstrap sample, multiple imputation will be implemented, producing 10 imputed datasets per bootstrap sample. Within each imputed dataset, analyses will be performed to obtain an estimate at each timepoint for all 5,000 datasets (i.e., 500 bootstrap samples * 10 imputed datasets). The combined estimate will correspond to the average of all 5,000

estimates and the 95% confidence interval will be derived from their distribution, using the 2.5th & 97.5th percentiles.

4.2. Multiple Imputation

OPERA data represent routine clinical care in the US. Therefore, the timing and frequency of follow-up visits and testing varies across patients and providers. Such variations may give rise to a missing data problem in analyses relying on outcomes measured at specific timepoints. While windows have been established around each timepoint of interest to increase the number of VL measurements available for analysis, it remains likely that VL will be missing for different PWH in different time windows.

Complete case analysis in which observations with missing data are omitted is appropriate when data are *missing completely at random* (MCAR; i.e., the missing values do not depend on observed or unobserved data).[33, 34] With any other type of missing data, despite being the simplest and most commonly used strategy, complete case analysis can lead to severe bias.[33, 34] In addition, complete case analysis reduces the analytical sample size, thus reducing the efficiency of estimations.[33]

Limitations of complete case analyses may be alleviated with imputation of missing values when data are *missing at random* (MAR; i.e., the missing values depend on observed data only).[35] Multiple imputation means that multiple values are imputed for each missing value, resulting in multiple complete datasets. Statistical analyses are then conducted on each imputed dataset, and the results are combined.[35] Multiple imputation relies on predictive models to predict the missing values based on the observed data. It is therefore important to include all variables that could help predict the missing values.[18]

Within each bootstrap sample, multiple imputation will be used to ensure that eligible OPERA participants can contribute to all timepoint analyses even if their routine clinical care resulted in VL being measured outside of the ranges defined in the protocol. Missing VLs in OPERA will be handled with multiple imputation implemented using chained equations to produce 10 imputed datasets.

4.3. Propensity score-weighting

PS-weighing uses the PS to up- and down-weight individuals so that the weighted-population resembles the target population.[36] PS-weighting is advantageous over PS-matching for two reasons. First, the weighting approach allows the utilization of all data available, whereas matching may require the exclusion of individuals who do not have an acceptable match. Matching may thus require a substantially larger sample size, which can be particularly problematic with smaller studies, such as this one.[37-39] Second, in some circumstances, PS-matching could be more biased: excluding individuals from the analysis may introduce selection bias, and the inability to find an exact match can result in residual

confounding.[20, 40] Using PS-weighting could thus improve statistical power and reduce the risk of bias in this external comparison study.

An external comparison study such as this one seeks to mimic the results of a hypothetical randomized controlled trial, had such a trial been conducted. Thus, the target population consists in the individuals who were recruited to participate in the trial (i.e., the treated group).[41] It is therefore important to ensure that the external control group resemble the treated group as closely as possible. The average treatment effect in the treated (ATT) can be estimated using standardized mortality ratio (SMR)-weighting. Of note, both SMR-weighing and PS-matching have been found to closely reproduce estimates from randomized controlled trials.[20]

SMR-weighting will thus be used to assess the ATT of IBA+OBR compared to non-IBA OBR, ensuring that the distribution of characteristics in the control group is the same as in the treated group.[20] In other words, the OPERA external control group will be weighted to look like the trial IBA+OBR-treated group. SMR-weighting relies on the PS, which corresponds to the probability for each person to receive IBA+OBR.

Within each imputed dataset, logistic regression will be used to obtain the PS and will include baseline values for important predictors of regimen selection. Continuous variables will be modeled flexibly with either quadratic terms or restricted cubic splines. Candidate covariates for the PS model include age, sex, VL, CD4 cell count, index date, agents in the OBR, resistance to ARVs, and the number of potent ARV in the OBR scaled by level of resistance. SMR weights are set to 1 in the treated population (i.e., IBA+OBR in TMB trials). In the control population (i.e., non-IBA OBR in OPERA), SMR weights are calculated as PS/(1-PS).[20] The final set of variables included in the PS model will be determined by the distribution of weights, as well as the balance of covariates across groups.

The distribution of weights will be evaluated for the presence of large weights. Model specification will be adjusted to obtain a mean weight as close to 1 as possible and avoid extreme weights. Weight trimming may be considered to reduce the impact of extreme weights.[42] In such case, trimming would be performed at the 1st and 99th percentile: all weights with a value below the 1st percentile (above the 99th percentile) would be set to the value of the first percentile (99th percentile).[42] The extent of trimming will be assessed based on the distribution of weights and adjusted as needed. The balance of covariates included in the PS model across groups will be assessed using standardized differences estimated before and after applying the SMR weights. A threshold of <10% will indicate an acceptable balance.[43]

4.4. Primary Objective: External Comparison of Virologic Effectiveness

The comparative virologic effectiveness of IBA+OBR vs. non-IBA OBR will be assessed at each timepoint for both viral suppression (VL < 200 copies/ml) and undetectability (VL < 50

copies/ml). Within each imputed dataset, estimates of risk will be derived from SMR-weighted Kaplan-Meier curves. The cumulative probability of viral suppression or undetectability will be estimated at each timepoint of interest for each arm, from which risk ratios will be calculated.

The distribution of estimates from all 5,000 datasets will be used to obtain the final estimate (mean) and 95% confidence intervals (2.5th and 97.5th percentiles).

4.5. Sensitivity Analyses

4.5.1. Impact of the viremia threshold in the OPERA non-IBA OBR group

In the TMB-202 and TMB-301/311 trials, the protocol restricted trial participation to individuals with a VL >1,000 copies/mL at baseline. However, in practice, individuals with a baseline VL ≤1,000 copies/mL were recruited. In routine clinical care, changes in regimen often occur before a threshold of VL >1,000 is reached. Therefore, to maximize the sample size and align with the population recruited in the trials, the main analysis will include PWH with a viral load > 200 copies/mL at the start of the index regimen in the OPERA non-IBA OBR group. However, to assess the impact of this more lenient threshold, a sensitivity analysis will be conducted in which the OPERA non-IBA OBR group will be restricted to PWH with a VL >1,000 copies/mL at index, as was intended for the trials.

Analyses described in sections 4.1-4.4 will be repeated after restricting the OPERA non-IBA OBR group to those with an index VL >1,000 copies/mL. New bootstrap samples and imputed datasets will be generated, within which SMR-weighted risk ratios will be estimated before being combined in a final estimate.

4.5.2. Replication of the European Registry Study

To ensure the comparability of results, analyses from the European registry Study will be repeated, using the R codes developed by FIECON. This analysis will use PS-matching; the covariates included in the PS will consist of baseline VL, baseline CD4 cell count, number of potent ARVs scaled by the level of resistance, and the calendar period score (lowest score = most comparable;

Table 3). A complete case analysis will be conducted. The subset of the study population identified for the main analysis with a viral load measurement at one or more timepoints of interest will be eligible for matching with replacement. Design and analyses will be adapted to take into consideration the differences between the European registry and the OPERA cohort.

Table 3. Relative comparability of regimens based on calendar periods

Calendar Period	Relative Comparability
Jan 2020 – Jan 2018	2
Jan 2018 – Jan 2016	1
Jan 2016 – Jan 2014	5
Jan 2014 – Jan 2012	3
Jan 2012 – Jan 2010	4
Jan 2010 – Jan 2008	6
Jan 2008 – Jan 2006	7
Jan 2006 – Jan 2004	8
Jan 2004 – Jan 2002	9
Jan 2002 – Jan 2000	10

5.Sample Size Considerations

IBA+OBR-treated group

In the TMB-202 trial, 113 patients were randomized to 800mg IBA every two weeks (n=59) or 2000mg IBA every four weeks (n=54). In the TMB-301/TMB-311 trials, 40 patients were enrolled and received 800mg IBA every 2 weeks following a 2000mg IBA loading dose.

The 59 TMB-202 participants randomized to 800mg IBA every two weeks and the 40 TMB-301/311 participants will be included, for a total sample size of 99 participants in the treatment group.

Non-IBA OBR control group

It is estimated that approximately 200 PWH in OPERA would be eligible for consideration in this study based on exposure to either DTG BID or DRV BID and presence resistance testing. If the final number of eligible PWH in OPERA is too small, the inclusion of more PWH on other regimen indicative of HTE will be considered.

6.Limitations, Quality & Ethical Considerations

6.1. Limitations of Proposed Methods

With approximately 14% of the HIV population that is linked to care in the OPERA database (per the CDC estimates), OPERA can provide detailed information on a large portion of the HIV population in the US. Even so, issues confronting population-level assessments include such aspects as differential medical care by practice size and specialty, academic and research orientation of the health care practitioner, ethnic-based and

gender-based attitudes and geographic regional health care practices. OPERA clinical data is collected at point-of-care and is subject to the record-keeping practices of each healthcare provider and the standards of each clinic or organization. Patients may see multiple physician practices for various conditions, which may result in incomplete case ascertainment. Data is collected for the medical management of patients and is not directly intended for research purposes, but rather for the care and management of individual patients and patient populations.

This study is an indirect comparison of data from two clinical trials and from the OPERA observational cohort. While most trial participants were in the US, differences in the distribution of key characteristics may still impact comparability of the groups. Moreover, in the main analysis, the sample size for the IBA+OBR group is limited to 99 participants on two different dosing schedules. Sensitivity analyses have been planned to assess the robustness of findings to different design choices.

6.2. Quality Control & Quality Assurance

Epividian has working practices and procedures governing the use of observational data, development of analysis specifications and plans, the development of analytical programming and the analytical quality assurance (QA) process and the scientific review of reports as well as clinical advisory charters for the clinical review of output intended for public domain. Working practices for the development of analysis specifications include basic identifying information, background material, relevant definitions of key study variables, population definitions, baseline definitions, specific requirements for dataset creation, and statistical requirements such as eligibility criteria, exposures, outcomes and model fitting. Working practices for programming include naming conventions, proper code documentation and commentary, content, appearance, efficiencies (i.e., use of macros), and organization of output, maintainability and generalizability. Working practices for programming QA include self-reviews of observational counts, missing data values, many-to-many merges, variable formatting, numeric-character & character-numeric conversions, uninitialized variables, unresolved macro references, report completeness and report-to-specification correspondence, and system errors and logs. The QA team review may include small sample spot-checking, coding log reviews, complete coding review, selected observations from intermediary dataset reviews, and/or independent programming to reproduce the results. Documentation of non-public domain reports includes market, scientific, statistical, and clinical review. Documentation of scientific protocols, reports and manuscripts intended for public domain follows two sequential steps: an internal-to-Epividian epidemiological, statistical, and clinical review, followed by a clinical/epidemiological external advisory board review.

All analytical data, coding algorithms, QA documentation and report outputs will be retained per Epividian standard practices.

6.3. Ethical Considerations

Clinical information is originally compiled into separate CHORUS databases for each clinic. This PHI is used in the creation of the CHORUS analytics and reporting used by each practice and its providers as part of Quality Improvement activities in an effort to improve care of patients. The data collection occurs via a secure and encrypted connection as part of Epividian's privacy and security policies and systems, which are routinely reviewed by a third-party privacy and security advisory organization.

Subsequently, the clinical data in each CHORUS database is de-identified and aggregated into the OPERA® Database following the guidelines of the Health Insurance Portability and Accountability Act (HIPAA) and the Health Information Technology for Economic and Clinical Health Act (HITECH).

Business Associate Agreements (BAAs) in place between Epividian and all medical practices govern, following the guidelines established in HIPAA and HITECH, the encryption, transportation, aggregation, de-identification and use of all clinical data in either the CHORUS reporting platform or the OPERA® Database. All medical practices are responsible for obtaining proper HIPAA consent for their patients. With BAAs in place and subsequent de-identification, a separate informed consent for each individual, in a non-interventional study is not required.

All clinical data in CHORUS is PHI and managed as such according to HIPAA, HITECH and relevant state regulations. The CHORUS portal, as a Quality Improvement activity, is accessed securely by clinic staff to view PHI for only those patients seen at the practice. All clinical data is subsequently de-identified as per HIPAA and HITECH in OPERA® with all reports submitted at the aggregated population level in OPERA®. No personally identifiable information is available in the OPERA® Database. The OPERA® Epidemiology & Clinical Advisory Board provides clinical and methodological review & oversight.

7. References

1. Laprise C, de Pokomandy A, Baril J-G, Dufresne S, Trottier H. Virologic Failure Following Persistent Low-level Viremia in a Cohort of HIV-Positive Patients: Results From 12 Years of Observation. *Clinical Infectious Diseases* 2013; 57(10):1489-1496.

2. WHO Expert Consultation. ART failure and strategies for switching ART regimens in the WHO European region. In: *Report of the WHO Expert Consultation*. Copenhagen.

3. Borroto-Esoda K, Waters JM, Bae AS, Harris JL, Hinkle JE, Quinn JB, et al. Baseline genotype as a predictor of virological failure to emtricitabine or stavudine in combination with didanosine and efavirenz. *AIDS research and human retroviruses* 2007; 23(8):988-995.

4. Al-Dakkak I, Patel S, McCann E, Gadkari A, Prajapati G, Maiese EM. The impact of specific HIV treatment-related adverse events on adherence to antiretroviral therapy: a systematic review and meta-analysis. *AIDS Care* 2013; 25(4):400-414.

5. Tang MW, Shafer RW. HIV-1 antiretroviral resistance: scientific principles and clinical applications. *Drugs* 2012; 72(9):e1-25.

6. Gardner EM, McLees MP, Steiner JF, Del Rio C, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. *Clin Infect Dis* 2011; 52(6):793-800.

7. Olson A, Bannert N, Sönnerborg A, de Mendoza C, Price M, Zangerle R, et al. Temporal trends of transmitted HIV drug resistance in a multinational seroconversion cohort. *AIDS* 2018; 32(2):161-169.

8. Rhee S-Y, Kassaye SG, Barrow G, Sundaramurthi JC, Jordan MR, Shafer RW. HIV-1 transmitted drug resistance surveillance: shifting trends in study design and prevalence estimates. *Journal of the International AIDS Society* 2020; 23(9):e25611.

9. Temesgen Z, Cainelli F, Poeschla EM, Vlahakis SA, Vento S. Approach to salvage antiretroviral therapy in heavily antiretroviral-experienced HIV-positive adults. *Lancet Infect Dis* 2006; 6(8):496-507.

10. Henegar C, Vannappagari V, Viswanathan S, DeKoven M, Clark A, Ackerman P, et al. Identifying heavily treatment-experienced patients in a large administrative claims database. In. International AIDS Society (IAS) Conference on HIV Science in Mexico City, Mexico, July 21-24, 2019.; 2019.

11. Hsu R, Henegar C, Fusco J, Vannappagari V, Llamoso C, Lackey P, et al. Identifying Heavily Treatment-Experienced Patients in the OPERA Cohort [Poster THPEB044]. In: *22nd AIDS Conference*. Amsterdam, the Netherlands; 2018.

12. Grover D, Copas A, Green H, Edwards SG, Dunn DT, Sabin C, et al. What is the risk of mortality following diagnosis of multidrug-resistant HIV-1? *J Antimicrob Chemother* 2008; 61(3):705-713.

13. Zaccarelli M, Tozzi V, Lorenzini P, Trotta MP, Forbici F, Visco-Comandini U, et al. Multiple drug class-wide resistance associated with poorer survival after treatment failure in a cohort of HIV-infected patients. *AIDS* 2005; 19(10):1081-1089.

14. Smith CJ, Phillips AN, Dauer B, Johnson MA, Lampe FC, Youle MS, et al. Factors associated with viral rebound among highly treatment-experienced HIV-positive patients who have achieved viral suppression. *HIV Med* 2009; 10(1):19-27.

15. Imaz A, Falco V, Ribera E. Antiretroviral salvage therapy for multiclass drug-resistant HIV-1-infected patients: from clinical trials to daily clinical practice. *AIDS Rev* 2011; 13(3):180-193.

16. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the Use of Antiretroviral Agents in Adults and Adolescents with HIV. In. Department of Health and Human Services; 2021.

17. Burger HU, Gerlinger C, Harbron C, Koch A, Posch M, Rochon J, et al. The use of external controls: To what extent can it currently be recommended? *Pharm Stat* 2021; 20(6):1002-1016.

18. Pugh SL, Brown PD, Enserro D. Missing repeated measures data in clinical trials. *Neurooncol Pract* 2022; 9(1):35-42.

19. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.

20. Kurth T, Walker AM, Glynn RJ, Chan KA, Gaziano JM, Berger K, et al. Results of Multivariable Logistic Regression, Propensity Matching, Propensity Adjustment, and Propensity-based Weighting under Conditions of Nonuniform Effect. *American Journal of Epidemiology* 2005; 163(3):262-270.

21. UNAIDS. UNAIDS Data 2021. In. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS); 2021.

22. Wandeler G, Johnson LF, Egger M. Trends in life expectancy of HIV-positive adults on antiretroviral therapy across the globe: comparisons with general population. *Current opinion in HIV and AIDS* 2016; 11(5):492-500.

23. Pelchen-Matthews A, Borges AH, Reekie J, Rasmussen LD, Wiese L, Weber J, et al. Prevalence and Outcomes for Heavily Treatment-Experienced Individuals Living With Human Immunodeficiency Virus in a European Cohort. *J Acquir Immune Defic Syndr* 2021; 87(2):806-817.

24. Lee JS, Cole SR, Richardson DB, Dittmer DP, Miller WC, Moore RD, et al. Incomplete viral suppression and mortality in HIV patients after antiretroviral therapy initiation. *Aids* 2017; 31(14):1989-1997.

25. Trickey A, May MT, Vehreschild J, Obel N, Gill MJ, Crane H, et al. Cause-Specific Mortality in HIV-Positive Patients Who Survived Ten Years after Starting Antiretroviral Therapy. *PloS one* 2016; 11(8):e0160460-e0160460.

26. Lalezari JP, Henry K, O'Hearn M, Montaner JS, Piliero PJ, Trottier B, et al. Enfuvirtide, an HIV-1 fusion inhibitor, for drug-resistant HIV infection in North and South America. *N Engl J Med* 2003; 348(22):2175-2185.

27. Khanna N, Klimkait T, Schiffer V, Irigoyen J, Telenti A, Hirschel B, et al. Salvage therapy with abacavir plus a non-nucleoside reverse transcriptase inhibitor and a protease inhibitor in heavily pre-treated HIV-1 infected patients. *AIDS* 2000; 14(7):791-799.

28. TaiMed Biologics Inc. Dose-Response Study of Ibalizumab (Monoclonal Antibody) Plus Optimized Background Regimen in Patients With HIV-1 (TMB-202) [NCT00784147]. In: ClinicalTrials.gov; 2014.

29. TaiMed Biologics Inc. Ibalizumab Plus Optimized Background Regimen in Patient With Multi-Drug Resistant HIV [NCT02475629]. In: ClinicalTrials.gov; 2020.

30. Castagna A. Italian Registry of HIV-1 Infected Patients With Drug-RESistant Virus to Reverse Transcriptase Inhibitors, InteGrasE and Viral Protease. (PRESTIGIO) [NCT04098315]. In: ClinicalTrials.gov; 2019.

31. Centers for Disease Control and Prevention. Diagnoses of HIV Infection in the United States and Dependent Areas, 2019. In: *HIV Surveillance Report*; 2021.

32. TaiMed Biologics Inc. Ibalizumab Plus Optimized Background Regimen in Treatment-Experienced Patients With Multi-Drug Resistant HIV-1 [NCT02707861]. In: ClinicalTrials.gov; 2021.

33. Harel O, Mitchell EM, Perkins NJ, Cole SR, Tchetgen Tchetgen EJ, Sun B, et al. Multiple Imputation for Incomplete Data in Epidemiologic Studies. *American Journal of Epidemiology* 2017; 187(3):576-584.

34. van Ginkel JR, Linting M, Rippe RCA, van der Voort A. Rebutting Existing Misconceptions About Multiple Imputation as a Method for Handling Missing Data. *Journal of Personality Assessment* 2020; 102(3):297-308.

35. Austin PC, White IR, Lee DS, van Buuren S. Missing Data in Clinical Research: A Tutorial on Multiple Imputation. *Canadian Journal of Cardiology* 2021; 37(9):1322-1331.

36. Webster-Clark M, Stürmer T, Wang T, Man K, Marinac-Dabic D, Rothman KJ, et al. Using propensity scores to estimate effects of treatment initiation decisions: State of the science. *Stat Med* 2021; 40(7):1718-1735.

37. Austin PC. The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. *Stat Med* 2014; 33(7):1242-1258.

38. Faria R, Hernandez Alava M, Manca A, Wailoo A. The use of observational data to inform estimates of treatment effectiveness in technology appraisal: methods for comparative individual patient data. NICE; 2015.

39. MA H, JM R. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC; 2020.

40. Rosenbaum PR, Rubin DB. The Bias Due to Incomplete Matching. *Biometrics* 1985; 41(1):103-116.

41. Sato T, Matsuyama Y. Marginal structural models as a tool for standardization. *Epidemiology* 2003; 14(6):680-686.

42. Lee BK, Lessler J, Stuart EA. Weight Trimming and Propensity Score Weighting. *PLOS ONE* 2011; 6(3):e18174.

43. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009; 28(25):3083-3107.