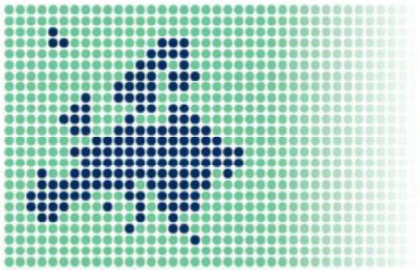


neatid The European treatment network for HIV, hepatitis and global infectious diseases

STATISTICAL ANALYSIS PLAN

Protocol Title:	**An open-label, multi-centre, randomised, switch study to evaluate the virological efficacy over 96 weeks of 2-drug therapy with DTG/RPV FDC in antiretroviral treatment-experienced HIV-1 infected subjects virologically suppressed with NNRTIs resistance mutation K103N**
SHORT TITLE/ACRONYM	WISARD:
SPONSOR NUMBER	SSCR105/NEAT33
CLINICAL PHASE	3b
EUDRACT NUMBER	2017-004040-38
SAP VERSION	1.0
SAP VERSION DATE	10.12.2020

Prepared by Study Analyst:

Name: Lambert ASSOUMOU

Date: 11-déc.-2020 | 08:17:54 PST

Signature: DocuSigned by: Lambert Assoumou E1CA74AEAFDD4D1...

DOCUMENT HISTORY

Document version number	Date	Reason for change

TABLE OF CONTENTS

1. Abbreviations and Definitions

Acronym	Description
ACR	Albumin: Creatinine Ratio
ACTG	AIDS Clinical Trial Group
AE	Adverse Event
ALT	Alanine Amino Transferase
ART	Anti Retroviral Therapy
ARV	Antiretroviral
AST	Aspartate aminotransferase
CD4	Cluster of Differentiation 4
CD8	Cluster of Differentiation 8
CKD-EPI	Chronic Kidney Disease Epidemiology
DAIDS	The Division of AIDS
DDI	Drug-Drug Interaction
DTG	Dolutegravir
ECG	Electrocardiogram
EFV	Efavirenz
EudraCT	European Clinical Trials Database
FDA	Food and Drug Administration (United States of America)
FDC	Fixed-Dose Combination
GFR	Glomerular filtration rate
HBsAb	Hepatitis B Surface Antibody
HBsAg	Hepatitis B Surface Antigen
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HDL	High Density Lipoprotein
HIV	Human Immunodeficiency Virus
HIV RNA	Human Immunodeficiency Virus Ribonucleic Acid
HS CRP	High-Sensitivity C-reactive protein
IMP	Investigational Medicinal Product
II	Integrase Inhibitors
INSTI	Integrase Strand Transfer Inhibitor
ISRCTN	International Standard Randomised Controlled Trials
IUD	Intrauterine Device
LDL	Low Density Lipoprotein
MCH	Mean Cell Haemoglobin
MCV	Mean Cell Volume
MDRD	Modified Diet in Renal Disease
NEAT ID	The European treatment network for HIV, hepatitis and global infectious diseases
NNRTI	Nonnucleoside Reverse Transcriptase Inhibitors

NNRTIs	Non-Nucleoside Reverse Transcriptase Inhibitors
NRTIs	Nucleoside Reverse Transcriptase Inhibitors
NVP	Nevirapine
OD	Overdose
PBMC	Peripheral Blood Monoclonal Cell
PCR	Protein: Creatinine Ratio
PI	Protease Inhibitor
PIs	Protease Inhibitors
PK	Pharmacokinetics
QoL	Quality of Life
QP	Qualified Person
RAL	Raltegravir
RBC	Red Blood Cell
RNA	Ribonucleic Acid
RPV	Rilpivirine
SAE	Serious Adverse Event
TDR	Drug Transmitted Resistance
ULN	Upper limit of Normal
VL	Viral Load
WBC	White Blood Cell
WOCBP	Women of Child Bearing Potential

2. Introduction

2.1 Preface

HIV-1 infected subjects that experience virological failure while on NNRTIs (EFV or NVP) with development of K103N mutation (+/- other NNRTIs and/or NRTIs mutations) are usually switched to a boosted PI-based regimen or other complex ARV combinations. The same is true for subjects who need to start antiretroviral therapy and have acquired virus that is already resistant to antiretrovirals (transmitted drug resistance, TDR). These "second line" combinations are often associated with numerous issues that can have a potential impact on the quality of life (QoL) of these patients: high pill burden, high potential for drug-drug interactions (DDIs), persistent side effects and metabolic toxicity. Therefore, a simpler and better tolerated alterative 2nd line treatment option would be a useful tool for the clinical management of these patients. Dolutegravir (DTG) is a well-tolerated 2nd generation integrase strand transfer inhibitor (INSTI) with a low metabolic impact and low DDI potential; rilpivirine (RPV) is a well-tolerated NNRTI not affected by the presence of the K103N mutation, characterized by a lower DDI potential and metabolic impact compared to other NNRTIs, neither requires a PK boosting agent.

2.2 Purpose of the analyses

This study will assess the efficacy and safety of switching from a boosted PI regimen to a dual combined therapy with DTG 50 mg OD + RPV 25 mg OD or continuing the current ART regimen in virologically suppressed patients with previous virological failure with NNRTIs, and having the clinically significant mutation K103N. The study will also assess whether a simplification of the treatment in term of pill burden, long term metabolic toxicity and potential of drug-drug interaction (DDI) improves the quality of life of the participants.

3. Study Objectives and Endpoints

3.1 Study Objectives

The primary objective of this trial is to compare viral load outcomes on DTG/RPV FDC regimen (experimental arm) to continued ART regimen (control arm) at 48 weeks as measured by HIV-1 RNA below or confirmed ≥50 copies/mL.

Secondary objectives are:

(a) To investigate whether switching patients to DTG/RPV FDC is associated with improvement of lipid profile, patient satisfaction, quality of life and potential for drug-drug interactions, investigated over time through 96 weeks with evaluation (and comparison to control arm) at week 24, 48 and 96.
(b) To evaluate DTG and RPV trough concentrations
(c) Changes in cell associated virus (using Peripheral Blood Monoclonal Cell (PBMC) analysis)

3.2 Study endpoints

3.2.1 Primary endpoint/outcome

The primary endpoint is the proportion of participants free of therapeutic failure at week 48 (HIV RNA <50 copies/mL and no discontinuation). Therapeutic failure is defined as virologic failure or discontinuation with HIV RNA <50 copies/mL. Discontinuation is defined by the study follow-up stop or any change or modification in the study treatment regimen.

The following procedure will be used to classify the participants.

1. Virological failure:

(a) Any patient with HIV RNA levels ≥50 copies/mL at any time points will have a repeat test on a second measurement within 2 weeks apart. If the result from the second measurement is <50 copies/mL, the patient will be categorized as HIV-RNA <50 copies/mL. If the result from the second measurement is ≥50 copies/mL, the patient will be categorized as virological failure.

(b) If a patient has HIV-RNA ≥50 copies/mL at a given time point and stops the study or at least one drug from the study treatment combination for any reason, he or she will be categorized as virological failure.

(c) If a patient discontinues the study or at least one drug from the study treatment combination at a given time point because of lack of efficacy then the patient will be categorized as virological failure.

2. Discontinuation (no virological data)

All discontinuations with HIV RNA <50 copies/mL will be considered as no virological data. There are three main reasons for no data at analysis time point:

(a) Discontinuation due to Adverse Event (AE) or Death.

Any patient who discontinues the study or the study treatment with HIV-RNA is < 50 copies/mL because of AE or death will be classified as discontinuation due to AE or death at from the date of discontinuation.

(b) Discontinuation for Other Reasons.

If a patient discontinues the study follow-up or changes or modifies any drug in the study treatment when his or her HIV-1-RNA is <50 copies/mL, he or she will be categorized as discontinued for other reasons at from the date of discontinuation.

(c) On study but missing data at analysis time point.

For example, if a participant has no data at analysis time point, but there is an HIV-RNA measurement after the analysis time point, this patient will be considered On Study but with Missing Data at analysis time point.

3.2.2 Secondary endpoints/outcomes

Secondary endpoints will include secondary efficacy endpoints, secondary safety endpoints, pharmacological endpoints, and secondary endpoints regarding patient related outcomes.

Secondary efficacy endpoints:

(a) Proportion of participants with virological failure (2 consecutive viral loads >50 copies at least 2 weeks apart or a viral load>50 copies/mL followed by a discontinuation of the study or the study treatment), investigated over 96 weeks.
(b) Incidence of genotypic resistance in case of confirmed HIV RNA>50 copies/mL and baseline factors associated with failure i.e. treatment emergent mutations (NRTI, NNRTI, PI and II regions) as compared with previous resistance tests.
(c) Proportion of participants with plasma HIV-1 RNA <50 c/mL at week 24, 48 and 96.
(d) Proportion of participants with virological suppression (<200 copies/ml HIV RNA) at week 24, 48, 96
(e) Changes from baseline in cell associated virus using PBMC Illumina MiSeq sequencing at week 48 and week 96.
(f) Changes from baseline in CD4, CD8 and CD4/CD8 ratio over 96 weeks with evaluation at week 24, 48 & 96.

Secondary safety endpoints:

(g) Number and percentage of participants with adverse events (AEs), severity of AEs, and treatment discontinuations due to AEs.
(h) Changes from baseline in laboratory parameter at week 24, 48 and 96.
(i) Changes from baseline in renal markers, bone markers, and fasting lipids at week 24, 48 and 96.

Secondary endpoints regarding pharmacological outcomes:

(j) Pre-dose plasma concentrations of DTG and RPV at week 4 (experimental arm only), 48 (experimental arm only) & 96. Plus pre/post-dose (as appropriate) at early termination visit and /or in the event of virological failure a post dose sample will be taken within 20-28 hours.
(k) Number of potential drug-drug interactions (DDIs) avoided (This will be measured by comparing the drug interaction outcomes between antiretroviral therapy and co-medications before and after the switch by using the www.hiv-druginteractions.org/ website (within the same study arm and between study arms). The grading will be done by looking at how many interactions graded amber or red (score of 1) will be graded green (score of 0) after the switch. For example, if a patient on atorvastatin as a co- medication and on two NRTIs and darunavir/cobicistat he would score "two amber and two green" and when switched to rilpivirine/dolutegravir, he would score "all green", showing the latter regimen is more favourable in terms of drug-drug interactions).

Interaction antiretroviral treatment at screening and atorvastatin

	Atorvastatin	score
NRTI 1	amber	1
NRTI 2	amber	1
Darunavir	green	0
Cobicistat	green	0
Total number of interaction (a)		2

Interaction between rilpivirine/dolutegravir and atorvastatin

	Atorvastatin	score
rilpivirine	green	0
doluteravir	green	0
Total number of interaction (b)		0

The number of potential DDIs avoided will be calculated as followed: a-b=2-0=2.

Secondary endpoints regarding patient related outcomes:

(l) Changes in quality of life from baseline.

The quality of life will be assessed by the EQ-5D-3L questionnaire in the present study. The questionnaire will be a self-administrated questionnaire at baseline, week 24, 48 and 96. The EQ-5D-3L descriptive system comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 3 levels: no problems (score of 1), some problems (score of 2), extreme problems (score of 3). And a health status score ranging from 0 to 100 (100 being worst quality of life) will be collected and will be considered as continuous variable.

(m) Changes in patient satisfaction from baseline.

The patients' satisfaction will be evaluated by self-reported questionnaires at baseline (HIVTSQs (status)) and post-baseline (HIVTSQc (change)) at weeks 24, 48 and 96. The questionnaires comprise 12 items and are shown in table below. The HIVTSQs questionnaire items are scored from 0 (dissatisfied) to 6 (satisfied). The HIVTSQc questionnaire items are scored from -3 (much less satisfied now) to 3 (much more satisfied now).

List of items used in HIVSTQs and HIVSTQc questionnaires.

ID	NAME
1	How satisfied are you with your current treatment?
2	How well controlled do you feel your HIV has been recently?
3	How satisfied are you with any side effects of your present treatment?
4	How satisfied are you with the demands made by your current treatment?
5	How convenient have you been finding your treatment to be recently?
6	How flexible have you been finding your treatment to be recently?
7	How satisfied are you with your understanding of your HIV?
8	How satisfied are you with the extent to which the treatment fits in with your lifestyle?
9	Would you recommend your present treatment to someone else who is being offered this HIV treatment?
10	How satisfied would you be to continue with your present form of treatment?
11	How easy or difficult have you been finding your treatment to be recently?
12	How satisfied are you with the amount of discomfort or pain involved with your present form of treatment?

(n) Changes in Pittsburgh sleep quality index with evaluation a week 24, 48, 72 and 96.

The Pittsburgh Sleep Quality Index (PSQI) contains 19 questions on the participant and 5 questions on the bed partner or roommate (if one is available). Only questions on the participant are included in the scoring. The 19-participant items are combined to form seven "component" scores, each of which has a range of 0-3 points. ln all cases, a score of "0"

indicates no difficulty, while a score of "3" indicates severe difficulty. The seven component scores are then added to yield one "global" score, with a range of 0-21 points, "0" indicating no difficulty and "21" indicating severe difficulties in all areas. The seven components are: component 1 (Subjective sleep quality); component 2 (Sleep latency); component 3 (Sleep duration); component 4 (Habitual sleep efficiency); component 5 (Step disturbances); component 6 (Use of sleeping medication); component 7 (Daytime dysfunction). More details on the derived variables are shown in the Appendix. Multiple imputation using Chained Equations approach (MICE) will be used to fill in missing data.

3.2.3 Exploratory endpoints/outcomes

(o) Changes in HS CRP over 96 weeks with evaluation at week 24, 48 and 96.

3.3 Derived variables

The following derived variables will be studied:

Variables	Calculation method
Age at inclusion, years	Difference between the date of enrolment and the birth date divided by 365.25
CD4/CD8 ratio	CD4 cell count divided by CD8 cell count
Duration on ART, years	Difference between the date of enrolment and the date of first antiretroviral treatment initiation divided by 365.25
Time since HIV diagnosis, years	Difference between the date of enrolment and the date of HIV diagnosis divided by 365.25
Duration of suppressed HIV RNA (<50 copies/mL), year	Difference between the date of enrolment and the date of last detectable viral load (HIV RNA >50 copies/mL) divided by 365.25
Changes from baseline in continuous biological parameters	Difference between value at each time point and the value at baseline
Percent change from baseline in continuous biological parameters	Difference between value at each time point and the value at baseline divided by the baseline value and multiply by 100
Number of potential DDIs avoided	Difference between number of interactions graded “amber or red” before switching and number of interactions graded “amber or red” after switching to dolutegravir/rilpivirine
Total cell associated HIV DNA copies per million PBMC	Total cell associated HIV DNA will be analysed as log transformed, as a continuous variable

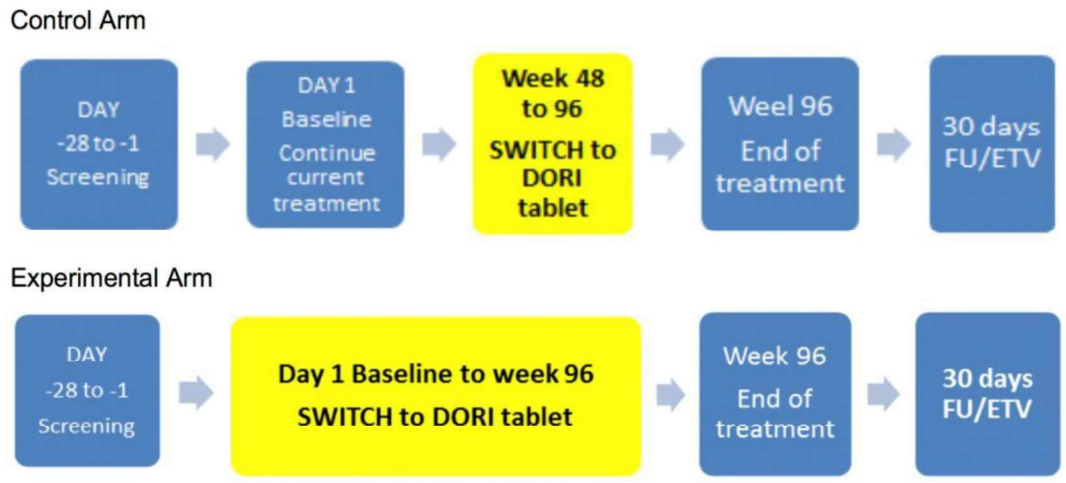
Partial dates will be extrapolated. For example, all dates saved as NK/FEB/2020 will be recoded as 15/FEB/2020 and those saved NK/NK/2020 will be recoded as 15/JUNE/2020. NK means not known.

4. Study Methods

4.1 General Study Design and Plan

This trial is an open-label, randomised, 2 arms, parallel-group, multicentre trial, comparing an immediate versus a deferred switch from a boosted PI containing regimen to a dual therapy with DTG and RPV in HIV infected patients with a history of virological failure with NNRTIs or baseline transmitted drug resistance and having developed a K103N mutation. The trial will be conducted within the NEAT ID network and include sites in 7 countries that have an excellent track record in clinical research. After screening, participants will be randomised 2:1 to switch to DTG/RPV at baseline or at 48 weeks.

The intervention will consist of immediate switch to DTG 50 mg OD + RPV 25 mg OD regimen at baseline or maintaining the current antiretroviral therapy for 48 weeks. At week 48, all participants in the deferred switch group with virological success will be switched to DTG 50 mg OD + RPV 25 mg OD therapy, and all participants will be followed for additional 48-weeks until week 96 (see figure).



After the screening visit, participants will be seen at baseline, weeks 4, 12, 24, 48, 52, 60, 72, 96, and either a follow-up visit (30 days (+/- 7 days) after week 96 visit) or an early termination visit. A week 60 visit may be an optional visit and performed for subjects on the control arm (3 months after therapeutic modification) as required per national guidelines. At each visit, participants will have routine investigations and will include viral load, CD4 & CD8, haematology (including haemoglobin, white cell count and differential, platelets), biochemistry (including sodium, potassium, creatinine, albumin, glucose, ALT, AST, ALP, total bilirubin, total cholesterol, HDL, LDL, triglycerides), PK, PBMC, quality of life questionnaires, patient satisfaction questionnaire, Pittsburgh sleep questionnaire, DDI assessment & urine sample (for haematuria, proteinuria, glycosuria, leukocytes) for full list of investigations please see section 4.4.

4.2 Randomisation Scheme and Blinding

4.2.1 Randomisation scheme

This study is open label and patients will be randomised to switch at baseline or after 48 weeks. Indeed, patients will be randomly allocated (2:1) to receive DTG/RPV at baseline in the experimental arm or to receive DTG/RPV after week 48 in the control arm. Randomisation will not be stratified by site because of the large number of sites relative to the sample size. The statistician of the trial will use central computer-generated permuted blocks of size 6 using a predefined randomised list to assign participants to treatment groups. Randomization will be done at Day 1 visit (date of the study treatment initiation in the experimental arm).

The participants in the experimental arm (Baseline visit switch group) will receive one combined Dolutegravir 50mg /Rilpivirine 25mg tablet taken orally once daily (at the same time) for up to 96 weeks with a meal. The participants in the control arm (deferred week 48 switch group) will continue their current boosted PI regimen (or other antiretroviral combination regimen) for 48 weeks. Patients will then be switched to one combined Dolutegravir 50mg /Rilpivirine 25mg tablet taken orally once daily for up to 48 weeks with a meal. Dolutegravir and Rilpivirine combined therapy tablets will be labelled and certified by Almac Group Ltd who will provide supply directly to each site on request of the sponsor. Boosted PIs (or other antiretroviral combination regimen) will be sourced from usual commercial supply as per standard practice.

This study is open label, therefore, all investigators, site pharmacists, study nurses and subjects will be unmasked and aware of the treatment allocation throughout the study.

4.2.2 Blinding

This section is not applicable

4.3 Study Variables

All the study variables are summarized in the following table.

Both treatment arms to complete all visits.	SCREENING	BASELINE (within 28 days of screening)	W4 (+/- 7)	W12 (+/- 7)	W24 (+/- 7)	W48 (+/- 7)	W52 (+/- 7)	W60 (+/- 7)	W72 (+/- 7)	W96 (+/- 7)	30 Day FU (+/- 7) (by phone)	ETV
Eligibility Criteria Assessment	x											
Informed Consent	x											
Demographic Data	x											

Medical and Social History Including • Recreational drug use • Smoking history • Alcohol intake • Concomitant diseases • Past & present medical history, including HIV-associated conditions • Antiretroviral history (including resistance history i.e. record all resistance history by major mutations in NRTI, NNRTI,	x											
PI and II regions)												
Alcohol intake, smoking habit and recreational drug use		x			x	x				x		x
Physical Examination (Including Height)	x											
Weight	x	x	x	x	x	x	x	x	x	x		x
Symptom Directed Physical Exam		x	x	x	x	x	x	x	x	x		x
Vital sign • Pulse • Blood pressure	x	x	x	x	x	x	x	x	x	x		x
Urine pregnancy test for women of child bearing potential (WOCBP) (β-HCG) only. Investigators to remind females of reproductive potential of the need to avoid pregnancy while in study and adherence to the study's contraception requirements. As a reminder, females of reproductive potential who changed their minds and desire to be pregnant should also be withdrawn from the study.	x	x (Predose)	x	x	x	x	x	x	x	x	x	x

Urinalysis (all tests should be performed on random spot urine. If random spot testing is not possible then 24hr urine can be performed, but only if pre-agreed with sponsor in writing) Including: • Protein:Creatinine Ratio (PCR)[2] • Albumin:Creatinine Ratio (ACR)[3] • urinary creatinine • urinary phosphate • urinary glucose (Can be performed quantitatively or by dipstick; optional if pre-agreed with sponsor in writing)	x	x			x	x				x		
beta-2 microglobulin (optional if pre- agreed with sponsor in writing)												
ECG	x											
CD4 & CD8 (count / % and ratio)	x	x			x	x			x	x		x
HBV/HCV (HBV antigen test (including HbsAG, anti-HBc and HBsAb) and HCV antibodies test (if test positive for HCV antibodies, test for HCV	x					x						
RNA, if positive and chronic for either HBV or HCV, add AST test to the biochemistry panel at screening visit. Also include Hep B surface antibodies post vax levels)												

Haematology Including: • Platelet count • RBC Count (indices MCV & MCH) • WBC count (absolute) • Haemoglobin • Hamatocrit RBC Count indices: • MCV • MCH Automated WBC differentials: • Neutrophils • Lymphocytes • Monocytes • Eosinophils Basophils	x	x	x	x	x	x	x	x	x	x		x
Clinical chemistry (Fasted for at least 10 hours except for screening) Including: • Urea • Chloride • Alkaline phosphatise • Creatine phosphokinase • Creatinine • Calcium • Phosphate • Glucose (optional in	x*	x*	x	x	x*	x*	x	x	x	x*		x
urine if agreed with sponsor in writing) • Total bilirubin • Potassium • HsCRP (optional if agreed with sponsor in writing) • Total protein • Sodium • ALT • Albumin • Creatinine clearance (Calculations: Cockcroft-Gault (requires weight), MDRD, CKD-Epi) *(* including urinary chemistry)*												

Fasting Lipids (Fasted for at least 10 hours except at screening) Including: • Total cholesterol • HDL cholesterol • LDL cholesterol • Triglycerides	x	x	x	x	x	x	x	x	x	x		x
HIV RNA viral load		x	x	x	x	x	x	x	x	x		x
Predose Pharmacokinetic assessment (A post-dose sample will be taken at early termination visit and/or in the event of virological failure within 20-28 hours of last dose)			x (Experimen tal arm only)			x (Experimen tal arm only)				x		x
Blood sampling for PBMC isolation (Required unless exemption documented in site contract i.e. if site does not have capacity to process PBMC samples)		x				x				x		x
Quality of Life Questionnaire		x			x	x				x		x
Patient Treatment Satisfaction Questionnaire		x			x	x				x		x
Pittsburgh Sleep Questionnaire		x			x	x				x		x
DDI Assessment		x			x	x				x		x
Adherence Assessment (Self report at visit) (including date and time of last dose on PK days,		x	x	x	x	x	x	x	x	x		x
including ETV visit and/or in the event of virological failure)												
AE Assessment		x	x	x	x	x	x	x	x	x	x	x
Con Med Check	x	x	x	x	x	x	x	x	x	x		x

5. Sample Size calculation

This is a pilot proof of concept study to investigate whether complex regimens taken by patients with resistance mutations can be replaced by a simple two class single pill regimen safely in terms of viral load and adverse events.

No formal sample size calculation is performed. It is anticipated that data from 150 participants is adequate to meet the trial objectives (100 in the DTG/RPV arm and 50 in the control arm). The expected success rate is 95% based on the data reported by the SWORD 1 & 2 studies (1). If the observed success rate is 95% in the DTG/RPV arm at week 48 and varies from 95% to 100% in the

control arm, with 150 participants (100 in the DTG/RPV arm and 50 in the control arm), the lower bound of the confidence interval of the difference in the success rate between the 2 arms (DTG/RPV – control) would varied from -9.3% to -7.4% (see table below).

Success rate			95% Confidence Interval	
DTG/RPV (n=100)	Control (n=50)	Difference	lower	upper
95%	95%	0.0%	-7.4%	7.4%
95%	96%	-1.0%	-7.9%	5.9%
95%	97%	-2.0%	-8.4%	4.4%
95%	98%	-3.0%	-8.8%	2.8%
95%	99%	-4.0%	-9.1%	1.1%
95%	100%	-5.0%	-9.3%	-0.1%

In addition, we would like to show whether the observed success rate in the DTG/RPV arm is above 95%. Therefore, we estimated power for a one-sample proportion test (Binomial test) by varying the external threshold from 80% to 90% to detect the minimum effect size at which we will have 80% power to achieve the goal (see table below).

Ho (null hypothesis): p = p0 versus Ha (alternative hypothesis): p > p0

N	**alpha**	**p**	**p0**	**delta**	**power**
100	5%	95%	80%	15%	100%
100	5%	95%	81%	14%	99.9%
100	5%	95%	82%	13%	99.6%
100	5%	95%	83%	12%	98.9%
100	5%	95%	84%	11%	97.2%
100	5%	95%	85%	10%	93.7%
100	5%	95%	86%	9%	93.7%
100	5%	95%	87%	8%	87.2%
100	5%	95%	88%	7%	76.6%
100	5%	95%	89%	6%	61.6%
100	5%	95%	90%	5%	43.6%

We show that, with 100 individuals in the DTG/RPV arm, we will have a 95% probability to discard a combination for which efficacy is smaller than 87% and we will select with a power of 80% the strategy for which the efficacy is above or equal to 95%.

The power calculation was made using the statistical software package nQuery Advanced, using the exact test for single proportion module (Version 8.5.2.0).

An interim analysis will be performed when 120 participants will reach the week 24 visit (corresponding to 80 participants in the DTG/RPV arm and 40 participants in the control arm). In the ANRS 163 ETRAL trial (Katlama et al., JAC 2019)(2), in which 10% of participants harboured viruses with K103N mutation, the proportion of participants in success at week 24 with a dual therapy with etravirine and raltegravir was estimated at 96.4% (95%CI 92.3 to 98.7). Therefore, for this interim analysis at week 24: The expected level of efficacy is 96%, and the unacceptable level of efficacy (p0) is 87%. With alpha = 5% (one-sided) and 80% power, if we include 80 participants in the DTG/RPV arm, we will consider that the treatment is effective to be continued if there are no more than 4 failures. We test the null hypothesis Ho: $p \leq 87\%$ versus the alternative hypothesis Ha: $p \geq 96\%$, with an 80% chance that the lower bound of the 95% confidence interval would exclude 87% if the true success rate (p) is 96%. This gives a 5% chance of continuing if p = p0 and an 80% chance of continuing if $p \geq 96$.

6. General Considerations

6.1 Timing of Analyses

An interim analysis will be performed to show whether DTG/RPV is able to maintain virological suppression at week 24, when 80% of the total participants will reach the 24-week visit (i.e. 120 participants: 80 in the DTG/RPV arm and 40 in the control arm) (anticipated to be November 2020).

The primary efficacy analysis will be performed at week 48, when all enrolled participants will reach the 48-week visit (anticipated to be May 2021).

The final analysis will be performed at week 96, when the last patient has completed his last scheduled visit in the protocol and when the database will be frozen (anticipated to be June 2022).

6.2 Timing of outcome assessments

Table A: Proposed Windows

Visit	Window (Through End-of-Study Week)	Window (Days)
12	6-18	43-126
24	18-30	127-210
48	42-54	295-378
72	66-78	463-546
96	90-102	631-714

6.3 Analysis Populations

6.3.1 Full Analysis Population

All randomly assigned participants will be included in the Full Analysis Population, except those who never received the study treatment. The participants who never received the study treatment will be excluded from the analysis. The primary endpoint and all secondary endpoints will be analysed on this Full Analysis population, called as a modified intention-to-treat (mITT) analysis.

6.3.2 Per-Protocol Population

The per-protocol population will include all patients from the mITT population, excluding participants with serious protocol violations which could interfere with the reliability of the efficacy outcome. Examples would include taking other antiretroviral treatment during the course of the study, incorrect randomisation or early discontinuation from the study for reasons unconnected to efficacy or safety. This analysis will be conducted to check the reliability of the results from the primary efficacy analysis.

6.3.3 Switch Included Population

In addition, a "Switch Included" analysis will be conducted, to check the outcomes of participants who discontinued randomised treatment for any reason. This analysis will follow the same methods as FDA Snapshot, but will include HIV RNA data from all participants with results available at Week 48, whether they are on or off randomised treatment. There will be no formal non inferiority outcome analysis as this is a pilot study.

6.3.4 Safety Population

The safety population will include all subjects included in the mITT population who received at least one dose of the study treatment

6.4 Covariates and Subgroups

6.4.1 Subgroup analysis

The sample size of 150 patients is too small to allow meaningful subgroup analyses. The two analyses described above (Per Protocol and Switch Included) will assess the reliability of the results from the primary efficacy analysis.

6.4.2 Adjusted analysis

The treatment effect will also be assessed adjusted for baseline variables which might influence the 48-week outcome.

Factors associated with the outcome will be tested using univariable logistic regression models. The following baseline variables will be tested: Age, gender, ethnic (black vs white), transmission group (MSM vs Other), Known duration since HIV diagnosis, duration of ART therapy, Positive Hep B antigen (yes vs no), Positive Hep C antibody (yes vs no), CD4 (above versus below 350 cells/µL), CD4/CD8 ratio, NRTI-backbone, PI received at baseline (yes vs no), body mass index, alcohol intake (yes vs no), current smoking (yes vs no), recreational drug use (yes vs no), diabetes (yes vs no), metabolic parameters (LDL-c, HDL-c, total cholesterol (TC), triglycerides and TC/HDL ratio), glucose, total protein, hsCRP, Creatinine clearance (e-GFR), presence of M184V mutation at baseline (yes vs no), number of PI mutation, number of NRTI mutation and number of allowed NNRTI mutation. Continuous variables will be modelled as continuous or as tertiles based on the AIC of the corresponding univariable model. Variables with univariable p-value <0.10 will be retained for the multivariable analysis.

All potential confounders that are known in the literature will also be included, such as the presence of M184V mutation at baseline and CD4 count at baseline (above versus below 350 cells/µL).

6.5 Missing data

In the primary efficacy analysis (FDA Snapshot), participants with missing data at analysis time points will be considered as having lost undetectable plasma HIV-1 RNA levels. The analyses of overall patient disposition will show the number of patients who did not attend each protocol-defined visit.

For all continuous efficacy and safety variables, we will use, in the mITT population, the LOCF (last observation carried forward) approach to deal with missing data. We will also use mixed model as a sensitivity analysis.

For patients reported outcomes, we will use multiple imputation approach to fill in missing data.

6.6 Interim analysis

One administrative statistical interim analysis is planned on the primary efficacy endpoint to show whether the efficacy of the dual therapy with DTG and RPV is above 96% after week 24. This interim analysis is planned to take place when 120 participants (80 in the DTG/RPV arm and 40 in the control) have reached the 24-week visit. It will be done following week 24 with the mITT population. The 95% two-sided confidence interval of the percentage of patients in therapeutic success will be calculated and DTG/RPV will be considered as an acceptable strategy if the lower bound is greater than 87%. The strategy will be considered unacceptable if there are 18 therapeutic failures or more. With 18 failures among 80 subjects in the DTG/RPV arm, the observed proportion of success is 77.5% and the associated CI is 66.7% to 86.1%, so just excluding 87%. Additional issues could include treatment

emergent drug resistance, serious adverse events or high rates of discontinuation for adverse events. The interim analysis will also explore whether the lower limit of the 95% confidence interval of the difference in the percentage of participants with therapeutic success between the DTG/RPV arm and the control arm (DTG/RPV- control) remain in acceptable range (above -13%).

6.7 Multi-centre Studies

Potential centre effects will not be taken into account in the analyses, as randomisation should make the two populations similar in all respects.

6.8 Multiple testing

As only an administrative interim analysis will be conducted, no multiplicity adjustment is required.

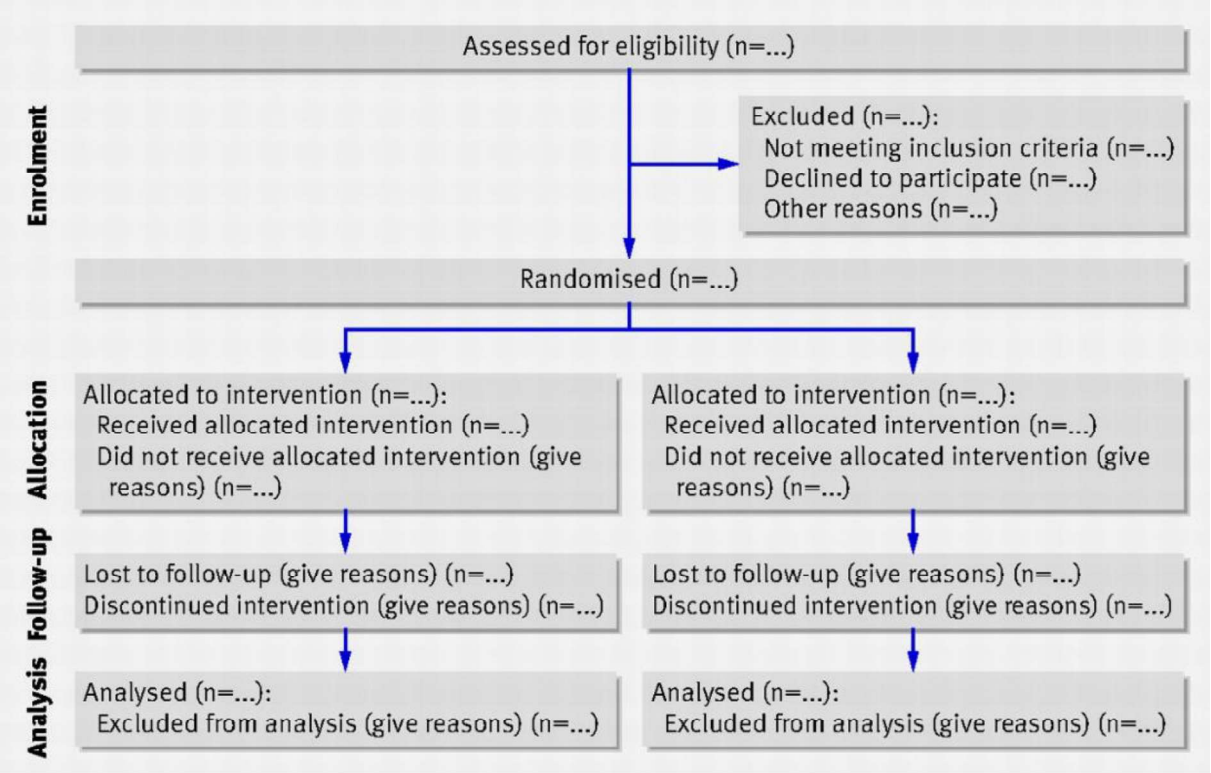
7. Summary of Study Data

All continuous variables will be summarized using the following descriptive statistics: n (non-missing sample size), mean, standard deviation, 1st quartile, median, 3rd quartile, maximum and minimum. The frequency and percentages (based on the non-missing sample size) of observed levels will be reported for all categorical variables.

7.1 Subject Disposition

The number of patients and the flowchart of the study will be presented. The period of enrolment and the total number of patients screened to be included in the study will be presented. The number of ineligible patients and the total number of randomised patients will be presented. The number of patients who never take the study treatment will also be presented by group and the reasons will be given and those who remained on the study treatment up to week 48 and 96 will be presented. The number of patients who discontinued the study will be presented and the reasons of discontinuation as well as the time from randomisation to discontinuation will be given. All these information will be shown in the study Flowchart as below (Figure 1).

Figure 1: Study Flowchart



7.2 Protocol Deviations

The following violations are defined as protocol deviations for the purpose of analysis: The deviation might be related to:

(a) Eligibility;
(b) Study related assessments;
(c) Social history;
(d) PK sampling;
(e) Drug dispensing;
(f) Adherence;
(g) Follow up (lost to follow-up, ….);
(h) Other (consent withdrawal, giving up, …)

The following violations will be considered as a serious deviation for the purpose of analysis: Patients who will not take study treatment; Interruptions of study drug for >10% of total time on treatment; Patient who will not fulfil the eligibility criteria; Patients who will withdraw the consent; Patients who will not comply with their randomization arm; Patients with missing value at the primary endpoint evaluation: Patient who will be lost to follow-up.

A patient will be considered lost to follow-up when he/she will not come to more than 2 visits of the

protocol. Patients who cannot be contacted on or before the last scheduled visit prior to the week 96 and who will not have a known reason for discontinuation (eg, withdrew consent, giving up, etc.) will be classified as “lost to follow-up”.

Patients will be considered in discontinuation of the study when they will refuse to continue the follow-up of the study.

7.3 Demographic and baseline variables

The baseline patients’ eligibility and characteristics of the mITT population will be described and will be presented in a summary table. This table will be structured with a column for each treatment group and a column for the total population in the order (DTG/RPV, control, and Total population). This summary table will be annotated with the total population size for each treatment group and the total population. The following variables will be described at baseline: Age, gender, ethnic, transmission group, child bearing potential, year of HIV diagnosis, Known duration since HIV diagnosis, duration of ART therapy, Positive Hep B antigen, Positive Hep C antibody, CD4 and CD8 cell count, CD4/CD8 ratio, antiretroviral treatment at baseline, NRT-backbone, PI received at baseline, body mass index, alcohol intake, smoking, recreational drug use, diabetes, metabolic parameters (LDL-c, HDL-c, total cholesterol (TC), triglycerides and TC/HDL ratio), glucose, total protein, hsCRP, Creatinine clearance (e-GFR), presence of M184V mutation, number of PI mutation, number of NRTI mutation and number of allowed NNRTI mutation. The summary statistics will be produced in accordance with section 7.

7.4 Concurrent Illnesses and Medical Conditions

The frequency of comorbidities will be described at baseline. The summary statistics will be produced in accordance with section 7.

7.5 Prior and Concurrent Medications

All medications ongoing at the start of the treatment or started during the study treatment will be described. The summary statistics will be produced in accordance with section 7.

7.6 Treatment Compliance

Adherence during the trial will be monitored by subject questioning regarding missed tablets at each visit. Compliance rate for each participant will be estimated by the number of pills consumed divided by the number of pills that will be theorically consumed and will be classified as low (<80%), medium (80-95%), and high (95-100%). The summary statistics will be produced in accordance with section 7. Generalised Estimating Equation (GEE) model will be used to compare the proportion of participants

with compliance rate <95% (not classified as high) between the 2 treatment groups over time with unstructured covariance matrix. This analysis will be performed if more than 20% of participants are not classified as high. The model will include treatment group, time and interaction between treatment group and time. Time will be considered as a categorical variable (0, 12, 24, 48, 60, 72, and 96 weeks) according to the planned visits. The proportion of participants with compliance rate <95% will be done at each visit.

8. Efficacy Analyses

8.1 Primary Efficacy Analysis

Virologic outcome will be determined by the last available measurement while the participant is on treatment and continued on trial within the time window (see Table A, section 6.2).

The primary efficacy analysis will be performed at week 48 with the mITT population. The proportion of patients with treatment success will be estimated by the number of patients with HIV-RNA<50 copies/mL up to week 48 with no discontinuation of the study or the study treatment at week 48 divided by the total number of participants in the mITT population. The associated two-sided 95% confidence interval will be calculated for each treatment group. We will also calculate the difference in treatment efficacy between the two treatment groups by measuring the difference in the percentage of participants with therapeutic success (i.e., DTG/RPV success rate minus the control success rate). The two-sided 95% confidence interval of the difference in therapeutic success rate will be given by the following formula (A):

$$(\widehat{p}_1 - \widehat{p}_2) \pm 1.96 \sqrt{\frac{\widehat{p}_1\widehat{q}_1}{n_1} + \frac{\widehat{p}_2\widehat{q}_2}{n_2}} \quad \text{where } \widehat{q} = 1 - \widehat{p}$$

The effect of study treatment will also be assessed in the mITT population using logistic regression model with treatment group adjusted for the presence of M184V mutation at baseline and CD4 count at baseline (above versus below 350 cells/µL) as well as all baseline variables potentially associated with the outcome, with univariable P-value <0.10 as described above in the section 6.4.2.

Sensitivity analyses will be performed with both the per-protocol and the switch included populations.

8.2 Secondary Efficacy Analyses

All secondary efficacy outcome analyses will be performed with the mITT population. All p-values will be two-sided with a significant level set at 0.05.

For all secondary efficacy outcomes the time window is shown in Table A (section 6.2).

8.2.1 Virological evaluation

(a) *Proportion of participants with virological failure (2 consecutive VL ≥50 copies/mL or one VL≥50 copies/mL and discontinuation of the study or the study treatment) over 96 weeks*

The Kaplan-Meier estimates will be used to determine the proportion of participants with virologic failure, investigated over the 96 weeks. The associated two-sided 95% confidence interval will be calculated for each group with Kalbfleisch and Prentice's formula. The Log-Rank test will be used to compare the virological failure rate between the 2 groups. An additional analysis using Cox regression model with treatment group adjusted for the presence of M184V mutation at baseline and CD4 count at baseline (above versus below 350 cells/µL) as well as all other potential confounders will be used to assess the effect of study treatment. We will also calculate the two-sided 95% confidence interval of the difference in the percentage of participants with virological failure (i.e., DTG/RPV failure rate minus the control failure rate) using the following formula (B):

$(\widehat{p_1} - \widehat{p_2}) \pm 1.96\sqrt{\widehat{s_1^2} - \widehat{s_2^2}}$ *where* $\hat{s}$ is the standard error of the proportion calculated with the Kaplan-Meier Method based on the Greenwood pointwise variance estimate.

The time from randomisation to virological failure is defined as the date of the first occurrence of virological failure or the date of last observation for participants without virological failure minus the date of initiating the study treatment. The date of virological failure is the date of the first of the 2 measurements conferring virological failure. The event will be the occurrence of virological failure. Participants in discontinuation for reasons other than virological failure will be censored at the date of last observation before evaluation.

(b) *Incidence of genotypic resistance in case of confirmed HIV RNA >50 and baseline factors associated with failure i.e. treatment emergent mutations (NRTI, NNRTI, PI and II regions) as compared with previous resistance tests.*

Genotypic resistance testing will be performed in participants with confirmed pVL >50 copies/mL and compared to previous resistance tests. Resistance interpretation will be given by using the Stanford algorithm (https://hivdb.stanford.edu/hivdb/by-mutations). The incidence of new resistance mutations will be described and compared between the two treatment groups. The frequency of new mutations conferring resistance to any antiretroviral drug will be given. A table will be built listing all participants with drug resistance mutations by treatment group. The table will include the following information: Patient Id, ARV treatment at baseline, age, sex, CD4 nadir, CD4 count at baseline, CD4/CD8 ratio at baseline, ARV treatment at failure, 1st VL at failure, 2nd VL at failure, resistance mutation at baseline, resistance mutation at failure, drugs resistant at baseline, drugs resistant at failure, drugs concentration at failure.

(c) *Proportion of participants with plasma HIV RNA <50 copies/mL at week 24, 48 and 96.*

The proportion of participants with plasma HIV RNA <50 copies/mL will be defined as the proportion of participants with HIV RNA <50 copies/mL and no discontinuation of the study or the study treatment. It will be estimated by the number of patients with HIV-RNA<50 copies/mL and no discontinuation at the analysis time point divided by the total number of patients in the mITT population. The associated two-sided 95% exact (Clopper–Pearson) confidence intervals will be calculated for each group as well as the two-sided 95% confidence interval of the difference in the proportion between the two treatment groups. The confidence intervals for the difference in proportions will be calculated with formula (A) defined above. Discontinuation is defined by the discontinuation of the study or any change or modification in the study treatment.

(d) Proportion of participants with plasma HIV RNA <200 copies/mL at week 24, 48 and 96.

The proportion of participants with plasma HIV RNA <200 copies/mL will be defined as the proportion of participants with HIV RNA <200 copies/mL and no discontinuation of the study or the study treatment. It will be estimated by the number of patients with HIV-RNA<200 copies/mL and no discontinuation at the analysis time point divided by the total number of patients in the mITT population. The associated two-sided 95% confidence interval will be calculated for each group as well as the two-sided 95% exact (Clopper–Pearson) confidence interval of the difference in the proportion between the two treatment groups. The confidence intervals for the difference in proportions will be calculated with formula (A) defined above. Discontinuation is defined by the discontinuation of the study or any change or modification in the study treatment.

(e) Change from baseline in cell associated virus using PBMC Illumina sequencing at week 48 and week 96

The evolution of cell associated HIV DNA (based on log transformation) will be described by boxplots for each treatment group at baseline, week 48 and 96. HIV DNA values below the limit of quantification (LOQ) value will be replaced by the LOQ value. The changes from baseline to week 48 and 96 will be compared between the two treatment group using mixed models for repeated measures with random intercept and unstructured covariance matrix, adjusted for baseline value. The models will include treatment group, time, interaction between treatment group and time, and baseline value as covariate. Time will be considered as categorical variable (0, 48, and 96 weeks). Visits window are shown in Table A (section 6.2). Data will be also summarised in tables by mean and standard error.

If the residual of the mixed model is not normally distributed and therefore does not allow the use of parametric test, the changes from baseline to week 48 and 96 of HIV DNA will be compared between the 2 groups using Mann-Whitney-Wilcoxon test.

8.2.2 Immunological evaluation

(f) Change from baseline in CD4, CD8, and CD4/CD8 ratio over 96 weeks with evaluation at week 24, 48, and 96.

The evolution of the CD4, CD8 and CD4/CD8 ratio over 96 weeks according to the treatment group will be described by boxplots for each treatment group at each visit. The changes will also be compared between the two treatment group using mixed models for repeated measures with random intercept and unstructured covariance matrix, adjusted for baseline value. The models will include treatment group, time, interaction between treatment group and time, and baseline value. Time will be considered as categorical variable (0, 12, 24, 48, 72 and 96 weeks) according the planned study visits. Visits window are shown in Table A (section 6.2). Data will be also summarised in tables by mean and standard error.

If the residual of the mixed model is not normally distributed and therefore does not allow the use of parametric test, the changes from baseline to week 24, 48 and 96 of CD4 T cell count, CD8 T cell count and CD4/CD8 ratio will be compared between the 2 groups using Mann-Whitney-Wilcoxon test.

9. Safety analyses:

(g) Number and percentage of participants with adverse events (AEs), severity of AEs, and treatment discontinuations due to AEs.

Adverse events (AE), especially grade 3 and 4 AE, serious adverse events (SAE), drug-related AE (all grades), treatment interruption due to AE (all grades), pregnancies and death will be summarized with descriptive statistics. The incidence rates of the events defined above will be compared between the treatment groups using a Poisson regression analysis which accounted for all adverse events and follow-up duration for each participant. The Poisson regression models will include the number of events, the treatment group and the duration of exposure in year in log. The number and percentage of death as well as the causes of death will be given. The number and percentage of pregnant women will also be given. A table will summarize the frequency of adverse events and pregnancies. This table will include the name of event, number and percentage of patients with events, number of events, and incidence rate (IR) per 100 person-years in the DTG/RPV group and the control group respectively, and incidence rate ratio (IRR) and 95%CI for IRR. All grades 3 or 4 AE related to the study drugs and treatment interruption due to AE will be put in the table.

(h) Changes from baseline in laboratory parameters at week 24, 48 and 96.

The evolution of haematology parameters (including haemoglobin, white cell count and differential, platelets) and biochemistry parameters (including sodium, glucose, ALT, AST, and total bilirubin) according to the treatment group will be described by boxplots for each treatment group at baseline, week 24, 48 and 96. The changes from baseline will be compared between the two treatment groups

using mixed models for repeated measures with random intercept and unstructured covariance matrix adjusted for baseline value. The model will include treatment group, time and interaction between treatment group and time, and baseline value. Time will be considered as categorical variables. Data will be also summarised in tables by mean and standard error.

If the residual of the mixed model is not normally distributed and therefore does not allow the use of parametric test, the non-parametric Mann-Whitney-Wilcoxon test will be used to compare the changes from baseline to week 24, 48 and 96 between the two treatment groups.

(i) Changes from baseline in renal markers, bone markers, and fasting lipids at week 24, 48 and 96.

The evolution of renal markers (including blood creatinine, blood creatinine clearance (eGFR), urine creatinine, urine protein, urine protein/creatinine ratio, blood albumin, urine albumin, urine albumin/creatinine ratio, urine beta 2 microglobulin, urine creatinine clearance (eGFR), potassium, blood phosphate, urine phosphate), bone markers (including ALP) and lipids fasting parameters (including total cholesterol, HDL, LDL, triglycerides) according to the treatment group will be described by boxplots for each treatment group at baseline, week 24, 48 and 96.

The changes from baseline will be compared between the two treatment groups using mixed models for repeated measures with random intercept and unstructured covariance matrix adjusted for baseline value. The model will include treatment group, time and interaction between treatment group and time, and baseline value. Time will be considered as categorical variables. Data will be also summarised in tables by mean and standard error.

If the residual of the mixed model is not normally distributed and therefore does not allow the use of parametric test, the non-parametric Mann-Whitney-Wilcoxon test will be used to compare the changes from baseline to week 24, 48 and 96 between the two treatment groups.

10.Pharmacokinetics

(j) Pharmacokinetic assessment

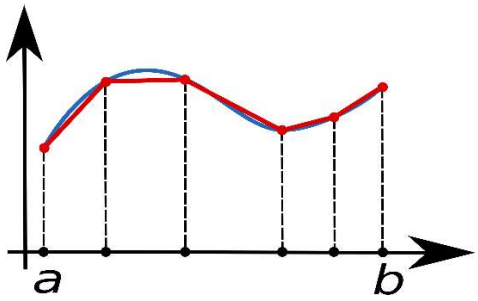
Pre-dose plasma concentrations of DTG and RPV at week 4 (experimental arm only), 48 (experimental arm only) and 96. Plus pre/post-dose (as appropriate) at early termination visit and /or in the event of virological failure a post dose sample will be taken within 20-28 hours.

At each time point, the AUC of DTG and RPV post administration will be summarized by the median, Q1, Q3, min, max and number of non-missing values. For each drug, we will compare the evolution of AUC over time using Wilcoxon paired test.

We will use the trapezoidal approach to calculate the area under the curve (AUC). For example, given data t0, t1, , tN and y0, y1, , yN , the area under the curve can be approximated as follows:

$$\int_a^b f(x)dx \approx \frac{1}{2}\sum_{i=1}^{N}(t_i - t_{i-1})(y_i + y_{i-1})$$

where t_i is the time value and y_i is the concentration of the antiretroviral drug.



(k) Number of potential DDIs avoided assessment

The number of potential DDIs avoided will be calculated by the difference between the number of interaction drugs graded "amber or red" before and after switching to dolutegravir and rilpivirine. More details are shown in section 3.2.2. The number of DDI avoided will be compared between the two-treatment groups using Mann-Whitney Wilcoxon test. Data will be summarized by the median, Q1, Q3, min, max and number of non-missing values.

11.Other analyses

(l) Changes in quality of life from baseline to week 24, 48 and 96.

The EQ-5D-3L quality of life questionnaires will be analysed using multiple imputation approach to replace missing values (assumption: Missing At Random). Multiple imputation using Chained Equations approach (MICE) will be used to fill in missing data. Therefore, we will create 10 datasets, in which missing data will be imputed from each patient's other covariables, including all items in the questionnaires, to obtain valid inference and reduce sampling variability resulting from the imputation process. The analysis will include all participants from the mITT population that will complete at least one of the 4 questionnaires (baseline, weeks, 24, 48, and 96). Participants with more than 20% missing data in a specific questionnaire will be excluded from the analysis of the questionnaire.
The changes from baseline will be compared between the two treatment groups using mixed models for repeated measures with random intercept and unstructured covariance matrix adjusted for baseline value. The model will include treatment group, time and interaction between treatment group and time, and baseline value. Time will be considered as categorical variables. Analyses will be run on each of the 10 datasets, including the imputed values, and the results will be combined with Rubin's rules. The p-values will be two-tailed, with a significant level of 0.05. Data will be summarised in tables by the mean and standard error.

To assess the impact of study treatment in the evolution of each of the 5 dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), a generalized estimating equation (GEE) model with a multinomial distribution (if 3 levels) or binomial distribution (if 2 levels) will be used. Each dimension: mobility, self-care, usual activities, pain/discomfort and anxiety/depression will be used as qualitative variables. If the numbers of reported level 3 problems are very low, we will group together level 2 and 3. Analyses will be run on each of the 10 datasets, including the imputed values, and the results will be combined with Rubin's rules. The model will include treatment group, time and interaction between time and treatment group, and baseline score as covariate. Time will be considered as categorical variable. The p-values will be two- tailed, with a significant level of 0.05.

(m) Changes in patient satisfaction from baseline.

The patients' satisfaction will be evaluated by self-reported questionnaires at baseline (HIVTSQs (status)) and post-baseline (HIVTSQc (change)) at weeks 24, 48 and 96. Each questionnaire include 12 items scored from 0 (dissatisfied) to 6 (satisfied) for HIVTSQs and scored from -3 (much less satisfied now) to 3 (much more satisfied now) for HIVTSQc. A global satisfaction score will be calculated at each time point by summing the result of the 12 items, leading to a global HIVTSQs score at baseline ranging from 0 to 72 (72 being higher satisfaction) and global HIVTSQc (change) post-baseline score ranging from -36 to 36 (36 being much more satisfaction). The analysis will include all participants from the mITT population that will complete at least one of the 4 questionnaires (baseline, weeks, 24, 48, and 96). We will compare the patients' satisfaction score post-baseline between the two treatment groups using mixed models for repeated measures with random intercepts and unstructured covariance matrix, adjusted for patients' baseline satisfaction score (HIVTSQs). The models will include treatment group, time and interaction between treatment group and time, and patients' baseline satisfaction score as covariate. Time will be considered as categorical variables. Data will be also summarised in tables by mean and standard error.

(n) Pittsburgh Sleep Quality Questionnaire

The Pittsburgh Sleep Quality Index (PSQI) contains 19 questions on the participant and 5 questions on the bed partner or roommate (if one is available). Only questions on the participant are included in the scoring. The 19-participant items are combined to form seven "component" scores, each of which has a range of 0-3 points. In all cases, a score of "0" indicates no difficulty, while a score of "3" indicates severe difficulty. The seven component scores are then added to yield one "global" score, with a range of 0-21 points, "0" indicating no difficulty and "21" indicating severe difficulties in all areas. The seven components are: component 1 (Subjective sleep quality); component 2 (Sleep latency); component 3 (Sleep duration); component 4 (Habitual sleep efficiency); component 5 (Step disturbances); component 6 (Use of sleeping medication); component 7 (Daytime dysfunction). Appendix shows how to calculate each component score.

As some data will be missing and because it is better to input than to ignore them, a multiple imputation using Chained Equations approach (MICE) will be used to fill in missing data. Ten imputations (M=10) will be chosen to obtain valid inference and reduce sampling variability resulting from the imputation process.

This analysis will include all participants from the mITT population that will complete at least one of the 4 questionnaires (baseline, weeks, 24, 48, 72, and 96). In each questionnaire, participants with more than 20% of missing data will be excluded from the analysis of the questionnaire.

The changes from baseline to week 24, 48, 72 and 96 will be compared between the 2 groups using mixed models for repeated measures with random intercept and unstructured covariance matrix, adjusted for patients' baseline score. The models will include treatment group, time and interaction between treatment group and time, and patients' baseline score as covariate. Time will be considered

as categorical variables. Analyses will be run on each of the 10 datasets, including the imputed values, and the results will be combined with Rubin's rules. The p-values will be two-tailed, with a significant level of 0.05. Data will be also summarised in tables by mean and standard error.

(o) Changes from baseline in hsCRP over 96 weeks with evaluation at week 24, 48 and 96

The evolution of hsCRP over 96 weeks according to the treatment group will be described by boxplots for each treatment group at baseline, weeks 24, 48 and 96.

The changes from baseline will be compared between the two treatment groups using mixed models for repeated measures with random intercept and unstructured covariance matrix adjusted for baseline value. The model will include treatment group, time and interaction between treatment group and time, and baseline value as covariate. Time will be considered as categorical variables. Data will be also summarised in tables by mean and standard error.

If the residual of the mixed model is not normally distributed and therefore does not allow the use of parametric test, the non-parametric Mann-Whitney-Wilcoxon test will be used to compare the changes from baseline to week 24, 48 and 96 between the two treatment groups.

12. Reporting Conventions

P-values ≥0.001 will be reported to 3 decimal places; p-values less than 0.001 will be reported as "<0.001". The mean, standard deviation, and any other statistics other than quantiles, will be reported to one decimal place greater than the original data. Quantiles, such as median, or minimum and maximum will use the same number of decimal places as the original data. Estimated parameters, not on the same scale as raw observations (e.g. regression coefficients) will be reported to 3 significant figures.

13. Technical Details

Data will be summarised using the IBM SPSS® statistics software version 24. The non-parametric analyses will be performed also with IBM SPSS statistics software version 24. The Poisson regression analysis, PK analysis and Figures will be done using Stata® SE software version 13. Generalised estimating equation and mixed models will be performed using SAS® software version 9.4.

14. References

1. Llibre JM, Hung CC, Brinson C, Castelli F, Girard PM, Kahl LP, Blair EA, Angelis K, Wynne B, Vandermeulen K, et al. Efficacy, safety, and tolerability of dolutegravir-rilpivirine for the maintenance of virological suppression in adults with HIV-1: phase 3, randomised, non-inferiority SWORD-1 and SWORD-2 studies. *Lancet.* 2018;391(10123):839-49.
2. Katlama C, Assoumou L, Valantin MA, Soulie C, Martinez E, Beniguel L, Bouchaud O, Raffi F, Molina JM, Fellahi S, et al. Dual therapy combining raltegravir with etravirine maintains a high level of viral suppression over 96 weeks in long-term experienced HIV-infected individuals over 45 years on a PI-based regimen: results from the Phase II ANRS 163 ETRAL study. *J Antimicrob Chemother.* 2019;74(9):2742-51.

APPENDIX: PITTSBURGH SLEEP QUALITY INDEX (PSQI)

INSTRUCTIONS: The following questions relate to your usual sleep habits during the past month only. Your answers should indicate the most accurate reply for the majority of days and nights in the past month. Please answer all questions.

1. What time did the subject usually go to bed at night? (24 hour clock)
   USUAL BED TIME __ ___ ___ ___ ___ _____ ___ _____

2. How long (in minutes) has it usually taken theb subject to fall asleep each night? NUMBER OF MINUTES___________

3. What time has the subject usually gotten up in the morning?
   USUAL GETTING UP TIME ________ _______ ________

4. How many hours of actual sleep did the subject get at night? {This may be different than the number of hours you spend in bed.)
   HOURS OF SLEEP PER NIGHT___________

INSTRUCTIONS: For each of the remaining questions, enter the one best response. Please answer all questions.

5. During the past month, how often has the subjject had trouble sleeping because they...

	Not during the past month	Lessthan once a week	Once or twice a week	Three or more times a week
(a) ...cannot get to sleep within 30 minutes	☐	☐	☐	☐
{b) ...wake up in the middle of the night or (b) wake up in the middle of the night or early morning	☐	☐	☐	☐
(c) ...have to get up to use the bathroom	☐	☐	☐	☐
{d ...cannot breathe comfortably	☐	☐	☐	☐
(e) ...cough or snore loudly	☐	☐	☐	☐
(f) ...feel tao cold	☐	☐	☐	☐
(g) ...feel tao hot	☐	☐	☐	☐
{h) ...had bad dreams	☐	☐	☐	☐
(i) ...have pain	☐	☐	☐	☐
(j) Other reason(s), please describe				
How often during the past month has the subject had trouble sleeping because of this (Other reason)?	☐	☐	☐	

	Very good	Fairly good	Fairly bad	very bad
6. How would the subject rate their sleep quality overall?	☐	☐	☐	☐

	Not during the past month	Less than once a week	Once or twice a week	Three or more times a week
7. How often has the subject taken medicine to help them sleep (prescribed or "over the counter")?	☐	☐	☐	☐
8. How often has the subject had trouble staying awake while driving, eating meals, or engaging in social activity?	☐	☐	☐	☐

	No problem at ail	Only a very slight problem	Somewhat of a problem	Avery big problem
9. How much of a problem has it been for the subject to keep up enough enthusiasm to get things done?	☐	☐	☐	☐

Partner / room mate information

	No bed partner or roommate	Partner/ roommate in other room	Partner in same room, but not same bed	Partner in same bed
10. Does the subject has a bed partner or room mate?	☐	☐	☐	☐

If the subject has a roommate or bed partner, how often in the past month would the bed partner or room mate say the subject experienced:

	Not during the past month	Less than once a week	Once or twice a week	Three or more times a week
(a) ...loud snoring	☐	☐	☐	☐
(b) ...long pauses between breaths while asleep	☐	☐	☐	☐
(c) ...legs twitching or jerking while you sleep	☐	☐	☐	☐
(d) ...episodes of disorientation or confusion during sleep	☐	☐	☐	☐

Has the subject experienced any other restlessness while sleeping?
Please describe

(e) How often during the past month has the subject experienced this other restlessness ?	☐	☐	☐	☐

SCORING INSTRUCTIONS FOR THE PITTSBURGH SLEEP QUALITY INDEX:

The Pittsburgh Sleep Quality Index (PSQI) contains 19 questions on the participant and 5 questions on the bed partner or roommate. Only questions on participnt are included in the scoring. The 19-participant items are combined to form seven "component" scores, each of which has a range of 0-3 points. In all cases, a score of "0" indicates no difficulty, while a score of "3" indicates severe difficulty. The seven component scores are then added to yield one "global" score, with a range of 0-21 points, "0" indicating no difficulty and "21" indicating severe difficulties in all areas.

Scoring proceeds as follows:

Component 1: Subjective sleep quality

Examine question #6, and assign scores as follows:

Response	Component 1 score
"Very good"	0
"Fairly good"	1
"Fairly bad"	2
"Very bad"	3

Component 1 score: _____

Component 2: Sleep latency

1. Examine question #2, and assign scores as follows:

Respo\nse	Score
:::15 minutes	0
16-30 minutes	1
31-60 minutes	2
> 60 minutes	3

Question #2 score: _____

2. Examine question #5a, and assign scores as follows:

Response	Score
Not during the past month	0
Less than once a week	1
Once or twice a week	2
Three or more times a week	3

Question #Sa score: _____

3. Add #2 score and #5a score

Sum of #2 and #Sa: _____

4. Assign component 2 score as follows:

Sum of #2 and #5a	Component 2 score
0	0
1-2	1
3-4	2
5-6	3

Component 2 score: _____

Component 3: Sleep duration

Examine question #4, and assign scores as follows:

Response	Component 3 score
> 7 hours	0
6-7 hours	1
5-6 hours	2
< 5 hours	3

Component 3 score:___ _

Component 4: Habitual sleep efficiency

1. Write the number of hours slept (question #4) here:____ _
2. Calculate the number of hours spent in bed:

   Getting up time (question #3):____ _

   Bedtime (question #1):____ _

   _____

   *Number of hours spent in bed:*____ _

3. Calculate habitual sleep efficiency as follows:

   (Number of hours slept/Number of hours spent in bed) X 100 = Habitual sleep efficiency (%)

   (_____/_____) X 100= %

4. Assign component 4 score as follows:

Habitual sleep efficiency %	Component 4 score
>85%	0
75-84%	1
65-74%	2
< 65%	3

Component 4 score:_ _

Component 5: Step disturbances

1 . Examine questions #5b-5j, and assign scores for each question as follows :

Response	Score
Not during the past month	0
Less than once a week	1
Once or twice a week	2
Three or more times a week	3

5b score: _____

5c score: _____

5d score: _____

5e score: _____

5f score : _____

5g score : _____

5h score: _____

5i score : _____

5j score: _____

2. Add the scores for questions #5b-5j:

Sum of #5b-5j: _____

3. Assign component 5 score as follows:

Sum of #5b-5j	Component 5 score
0	0
1-9	1
10-18-4	2
19-27	3

Component 5 score: _____ _

Component 6: Use of sleeping medication

Examine question #7 and assign scores as follows:

Response	Component 6 score
Not during the past month	0
Less than once a week	1
Once or twice a week	2
Three or more times a week	3

Component 6 score: _ _

Component 7: Daytime dysfunction

1. Examine question #8, and assign scores as follows:

Response	Score
Never	0
Once or twice	1
Once or twice each week	2
Three or more times each week	3

Question#B score: _____

2. Examine question #9, and assign scores as follows:

Response	Score
No problem at all	0
Only a *very* slight problem	1
Somewhat of a problem	2
A *very* big problem	3

Question #9 score : _____

3. Add the scores for question #8 and #9:

Sum of #8 and #9: _____

4. Assign component 7 score as follows:

Sum of #8 and #9	Component 7 score
0	0
1-2	1
3-4	2
5-6	3

*Component 7 score:*____ _

Global PSQI Score

Add the *seven* component scores together:

*Global PSOI Score:*____ _