Durability and tolerability of first-line regimens including two nucleoside reverse transcriptase inhibitors and raltegravir or ritonavir boosted-atazanavir or -darunavir: data from the ICONA Cohort

Antonella d'Arminio Monforte[1], MD; Patrizia Lorenzini[2], Stat; Alessandro Cozzi-Lepri[3], PhD, Stat; Cristina Mussini[4] MD; Antonella Castagna[5], MD; Franco Baldelli[6], MD; Massimo Puoti[7], MD; Francesca Vichi[8], MD; Adelaide Maddaloni[9], MD; Sergio Lo Caputo[10], MD; Nicola Gianotti[5], MD; Andrea Antinori[2] ,MD; on behalf of the Icona Foundation Study Group*

[1]Clinic of Infectious and Tropical Diseases, Department of Health Sciences, ASST Santi Paolo e Carlo, University of Milan. Milan, Italy; [2]HIV/AIDS Clinical Department, National Institute for Infectious Diseases "Lazzaro Spallanzani". Rome, Italy; [3]Department of Infection and Population Health, Division of Population Health, UCL Medical School, Royal Free Campus. London, United Kingdom; [4]Infectious Disease Clinic, University of Modena and Reggio Emilia. Modena, Italy; [5]Department of Infectious Diseases; San Raffaele Scientific Institute, University Vita-Salute San Raffaele. Milan, Italy; [6]Unit of Infectious Diseases, University of Perugia. Perugia, Italy; [7]Department of Infectious Diseases; ASST Grande Ospedale Metropolitano Niguarda. Milan, Italy; [8]Unit of Infectious Diseases, Santa Maria Annunziata Hospital. Florence, Italy; [9]Unit of Infectious Diseases in Women, Cotugno Hospital. Naples, Italy; [10]UOC of Infectious Diseases; Policlinico di Bari. Bari, Italy.

*The complete membership of the author group can be found in the Acknowledgments.

Corresponding author: Antonella d'Arminio Monforte (antonella.darminio@unimi.it)

Short Title: Durability of first line regimens including RAL or ATV/r or DRV/r

Partly presented at the HIV Glasgow 2016 Conference, 23-26 October 2016, abstract n P021

Acknowledgements

ICONA FOUNDATION STUDY GROUP:

BOARD OF DIRECTORS

A d'Arminio Monforte (President), A Antinori, A Castagna, F Castelli, R Cauda, G Di Perri, M Galli, R Iardino, G Ippolito, A Lazzarin, GC Marchetti, CF Perno, G Rezza, F von Schloesser, P Viale.

SCIENTIFIC SECRETARY

A d'Arminio Monforte, A Antinori, A Castagna, F Ceccherini-Silberstein, A Cozzi-Lepri, E Girardi, S Lo Caputo, C Mussini, CF Perno, M Puoti.

STEERING COMMITTEE

M Andreoni, A Ammassari, A Antinori, C Balotta, A Bandera, P Bonfanti, S Bonora, M Borderi, A Calcagno, L Calza, A Capetti, MR Capobianchi, A Castagna, F Ceccherini-Silberstein, A Cingolani, P Cinque, A Cozzi-Lepri, A d'Arminio Monforte, A De Luca, A Di Biagio, E Girardi, N Gianotti, A Gori, G Guaraldi, G Lapadula, M Lichtner, S Lo Caputo, G Madeddu, F Maggiolo, G Marchetti, S Marcotullio, L Monno, C Mussini, S Nozza, M Puoti, E Quiros Roldan, R Rossotti, S Rusconi, MM Santoro, A Saracino, M Zaccarelli.

STATISTICAL AND MONITORING TEAM

A Cozzi-Lepri, I Fanti, L Galli, P Lorenzini, A Rodano, M Shanyinde, A Tavelli

BIOLOGICAL BANK INMI

F Carletti, S Carrara, A Di Caro, S Graziano, F Petrone, G Prota, S Quartu, S Truffa

PARTICIPATING PHYSICIANS AND CENTERS

Italy A Giacometti, A Costantini, V Barocci (Ancona); G Angarano, L Monno, C Santoro (Bari); F Maggiolo, C Suardi (Bergamo); P Viale, V Donati, G Verucchi (Bologna); F Castelli, C Minardi, E Quiros Roldan (Brescia); T Quirino, C Abeli (Busto Arsizio); PE Manconi, P Piano (Cagliari); B Cacopardo, B Celesia (Catania); J Vecchiet, K Falasca (Chieti); L Sighinolfi, D Segala (Ferrara); P Blanc, F Vichi (Firenze); G Cassola, C Viscoli, A Alessandrini, N Bobbio, G Mazzarello (Genova); M Lichtner, I Pozzetto (Latina); P Bonfanti, I Caramma (Lecco); A Chiodera, P Milini (Macerata);G Nunnari (Messina); A d'Arminio Monforte, M Galli, A Lazzarin, G Rizzardini, M Puoti, A Castagna, G Marchetti, MC Moioli, R Piolini, AL Ridolfo, S Salpietro, C Tincati, (Milano); C Mussini, C Puzzolante (Modena); A Gori, G Lapadula (Monza); A Chirianni, G Borgia, V Esposito, R Orlando, G Bonadies, F Di Martino, I Gentile, L Maddaloni (Napoli); AM Cattelan, S Marinello (Padova); A Cascio, C Colomba (Palermo); F Baldelli, E Schiaroli (Perugia); G Parruti, F Sozio (Pescara); G

Magnani, MA Ursitti (Reggio Emilia); M Andreoni, A Antinori, R Cauda, A Cristaudo, V Vullo, R Acinapura, G Baldin, M Capozzi, S Cicalini, A Cingolani, L Fontanelli Sulekova, G Iaiani, A Latini, I Mastrorosa, MM Plazzi, S Savinelli, A Vergori (Roma); M Cecchetto, F Viviani (Rovigo); G Madeddu, P Bagella (Sassari); A De Luca, B Rossetti (Siena); A Franco, R Fontana Del Vecchio (Siracusa); D Francisci, C Di Giuli (Terni); P Caramello, G Di Perri, S Bonora, GC Orofino, M Sciandra (Torino); M Bassetti, A Londero (Udine); G Pellizzer, V Manfrin (Vicenza); G Starnini, A Ialungo(Viterbo).

no tracjs

Durability and tolerability of first-line regimens including two nucleoside reverse transcriptase inhibitors and raltegravir or ritonavir boosted-atazanavir or -darunavir: data from the ICONA Cohort

Abstract

Background: We aimed to mimic the ACTG 5257 trial, comparing raltegravir (RAL), ritonavir-boosted atazavavir (ATV/r) and ritonavir-boosted darunavir (DRV/r) in the observational setting.

Methods: All the ICONA patients starting a first cART with 2NRTI +ATV/r, DRV/r or RAL were included. Primary end-point was treatment failure, ie virological failure (confirmed HIV-RNA>200copies/ml >6 months therapy) or discontinuation for any reason of the third drug. Secondary end-points: virological failure50 (50 copies/mL threshold), and discontinuation of the third drug due to intolerance/toxicity. Cox regression analyses were run to compare the risk of outcomes between the three regimens.

Results: 2,249 patients were included, 985 (44%) initiated ATV/r, 1,023 (45%) DRV/r and 241 (11%) RAL; median follow-up of 3.6 years (IQR: 2.3-5.2). After controlling for baseline confounding factors, patients given ATV/r showed a 26% higher risk of treatment failure (TF) vs DRV/r (AHR 1.26, 95%CI 1.11-1.43); patients on RAL had a lower risk of TF vs ATV/r (AHR 0.81, 95%CI 0.66-0.99). The probability of virological failure50 was significantly lower for people initiating RAL vs DRV/r (AHR 0.46, 95%CI 0.24-0.87) or ATV/r (AHR 0.52, 95%CI 0.27-0.99). In addition, RAL was associated to a lower risk of discontinuation for toxicity vs both DRV/r (AHR: 0.37, 95%CI: 0.19-0.72) and ATV/r (AHR: 0.18, 95%CI: 0.09-0.34). ATV/r was associated with a higher risk of discontinuing due to toxicity (AHR 2.09, 95%CI 1.63-2.67) vs DRV/r.

Conclusions: In our observational study, we confirmed higher risk of treatment failure and lower tolerability of ATV/r-based regimens as compared to those including DRV/r or RAL.

Keywords: cohort study; antiretroviral regimens; therapy discontinuation; raltegravir;

Boosted-atazanavir; boosted-darunavir.

Introduction

Although newer drugs belonging to the integrase inhibitors class (raltegravir, dolutegravir and elvitegravir) as well as newer generation non-nucleoside reverse transcriptase inhibitors (NNRTI) (such as rilpivirine) are now the most commonly prescribed third agents in first-line combination antiretroviral therapy (cART), darunavir/r (DRV/r) and atazanavir/r (ATV/r) are still among the indicated alternative options in several treatment guidelines [1-3]. Indeed, ritonavir-boosted protease inhibitors (PI/r)-containing regimens retain strong supporting evidence of long-term clinical efficacy, and are still considered as first-line options in persons with low adherence or in cases with missing drug resistance tests before starting cART, due to their high genetic barrier [1-3].

The ACTG 5257 trial has compared the efficacy and tolerability of three first-line regimens including ATV/r, DRV/r or raltegravir (RAL), in combination with tenofovir/emtricitabine (TDF/FTC) in 1,809 naïve subjects enrolled in clinical sites in the United States [4]. The trial demonstrated similar virological potency of the three regimens, even in patients starting cART at high viral load, and lower tolerability for ATV/r including regimens as compared to the other two drugs and also lower tolerability for DRV/r as compared to RAL.

One limitation of the ACTG study is its open-label design, and people on ATV/r may have been more prone to switch their regimen for elevate bilirubin levels or the fear of a sustained elevation. Moreover, ACTG 5257 showed results up to 3 years from the date of regimens initiation and longer terms estimates are currently lacking.

We therefore aimed to conduct an analysis similar to that of the ACTG 5257 trial, by comparing the long-term durability and safety of first-line RAL-including regimens to therapies including either DRV/r or ATV/r but using observational data. Our analysis also provides a comparison of the effectiveness of the regimens when used in HIV-infected persons seen in routine clinical practice in Italy where, unlike the USA, there is no barrier to access to treatment and care.

Methods

The ICONA Foundation Study

The Italian Cohort Naives Antinetrovirals (ICONA) Foundation Study is a multi-centre observational study of HIV-1-infected patients set up in 1997, including 51 centres of Infectious Diseases across Italy. Patients eligible to be included in the cohort are those starting cART when they are naive to antiretrovirals, regardless of the reason for which they had never been previously treated. Demographic (age, sex, risk factors for HIV, education, job, marital status), clinical (all clinical events, both HIV and non HIV related) and laboratory data and information on therapy (both HIV and non HIV) are collected and recorded using electronic data collection and updated at any new event or at least twice a year [www.icona.org]. Details of the study are described elsewhere [5].

The ICONA Foundation study has been approved by IRB of all the participating centres. All patients sign a consent form to participate in ICONA, in accordance with the ethical standards of the committee on human experimentation and the Helsinki Declaration (1983 revision). The estimated percentage of refusal to participate the study is 5-10%.

Patient population

All the patients from the ICONA Foundation cohort who started their first cART regimen after January 1, 2008 (year in which RAL was licenced for use in Italy) with 2NRTI (either TDF+FTC or abacavir+lamivudine -ABC+3TC) + ATV/r or DRV/r or RAL were included in this analysis. We recorded the presence of comorbidities at ART initiation (baseline), defined as: any non AIDS-defining malignancy; cardiovascular events (acute myocardial infarction, coronary disease requiring invasive procedures, stroke); hepatic events (decompensated cirrhosis, i.e. variceal bleeding, porto-systemic encephalopathy, refractory ascites); kidney injury (onset of a confirmed estimated glomerular filtrate rate [eGFR] <60 ml/min using Modification of Diet in Renal Disease -MDRD- formula or kidney failure requiring dialysis or transplantation). All causes of discontinuation are collected in the ICONA database as reported by the treating physicians who are asked to indicate which the main reason for stopping was. Reasons include simplification (defined either as the reduction of number of drugs or the decrease in daily doses or pills), intolerance, toxicity, failure (virological, immunological or clinical), non-adherence, planned interruption (including end of pregnancy and medical decision) and other causes (patients decision, pregnancy, enrolment or ending of a clinical trial and drug-drug interaction).

Study outcomes

The response to the initial regimens was compared according to the specific third drug started with respect of a number of end-points. Our primary objective was to compare treatment failure between the three regimens (RAL, DRV/r, ATV/r). The composite end-point of treatment failure was defined as virological failure (confirmed HIV-RNA>200 copies/ml after 6 months of therapy) or discontinuation of the third drug of the regimen for any reasons. Secondary end-pointes included:

- virological failure 50: confirmed HIV-RNA >50 copies/mL after 6 months of therapy
- discontinuation of DRV/r or ATV/r or RAL because of intolerance/toxicity.

Discontinuations of the NRTI backbones have been ignored in this analysis.

Mean CD4 change from baseline to 2^{nd} years of follow-up-according to the third drug were also analysed in a subset of the study population with complete CD4 count data.

Patients were followed up from date of starting one of the studied regimens (i.e. baseline) to the first end-point event, November 15^{th}, 2017, death or loss to follow-up.

Statistical analyses

For the comparison of characteristics at time of treatment initiation among the three groups, Chi-square or Kruskal-Wallis test were used as appropriate. Survival analysis with Kaplan-Meier curves were used and the probability of the outcome was estimated together with 95% confidence interval for each time point. Log-rank test was used to test the equality of survival curves.

Cox regression analysis stratified by clinical site was employed to compare the risk of primary and secondary outcomes by means of computing unadjusted and adjusted (after controlling for potential measured confounding factors) hazard ratios. The proportional-hazards assumption was verified testing the interaction between the predictors and natural logarithm of survival time. All variables considered in the univariable model have been also included in the multivariable model. The adjusted analysis included the following a priori chosen, time-fixed covariates at cART initiation: age, gender (M, F), nation of birth (native, migrant), (Heterosexual, intravenous drug addicts-IDU-, men sex with men –MSM-, Other/unknown), hepatitis status (HCV-Ab+, HCVAb-, HBsAg+, HBsAg-, unknown), AIDS (yes no), (0-200 201-350, 351-500, 500+) and viral load (<20.000, 20.000-100.000, 100.000-250.000, 250.000+) and year of starting cART (2008-09, 2010-11, 2012-13, 2014-15), nucleoside pair (TDF/FTC, ABC/3TC) and third drug started (DRV/r, ATV/r, RAL). The reference group was also changed to allow a three-way comparison between RAL, DRV/r and ATV/r.

We have used a cause-specific hazards for the survival analysis. This was done under the non-testable assumption that censoring due to virological failure is non informative (unrelated to) for the risk of stopping a drug because of other reasons (e.g. toxicity or simplification).

Incidence rate of each endpoint was calculated as number of events over person-years follow-up (PYFU).

Patients with CD4 count at pre-cART and at 24 months (+/- 4 months) were selected and compared with subjects without this information. To define if the immunological recovery was different among the 3 regimens, univariable and multivariable linear regression was used. The following time-fixed covariates at cART initiation were considered: age, gender, nation of birth, mode of HIV transmission, hepatitis status, AIDS, CD4 count and viral load and year of starting cART, nucleoside pair and third drug started.

Results

<u>Characteristics of the Study Population</u>

A total of 2,249 patients fulfilling the criteria of inclusion were studied: 985 (43.8%) initiated a first ART regimen including ATV/r, 1,023 (45.5%) DRV/r and 241 (10.7%) RAL. The median age at baseline was 40 years (IQR: 32-48), 21% were females, 22% migrants, 40% men who acquired HIV through sex with other men (MSM); 224 (10%) were HCV coinfected and 92 (4.1%) HBV coinfected. Median CD4 at treatment initiation was 277 cells/mmc (IQR: 120-415), the proportion of subjects with baseline CD4 <200 was 37%. Median HIV-RNA at baseline was 4.9 log10 copies/mL (IQR: 4.3-5.4), 44% had a pre-treatment HIV-RNA >100,000 copies/mL.

Patients on ATV/r- were less frequently males, less frequently Italian, more frequently HCV coinfected and started cART in earlier calendar years than patients given either DRV/r or RAL. Patients on DRV/r had the lowest median CD4 counts and highest median HIV-RNA copy levels. Patients on RAL including regimens were more frequently affected by comorbidities (24/241; 10%) than those initiating ATV/r (42/985; 4.3%) or DRV/r (52/1023; 5.1%) (p=.002). Patients' characteristics according to the third drug are shown in Table 1.

Participants have been followed-up for a median of 3.6 years from ART initiation (interquartile range-IQR: 2.3-5.2) (ATV/r: 4.3, IQR: 2.7-5.7; DRV/r: 3.4, IQR: 2.3-4.9; RAL: 2.3, IQR: 1.5-3.5).

<u>Incidence rates of various endpoints</u>

Over 5,431 person-years of follow-up (PYFU), 1,433 patients reached the composite end-point of treatment failure, resulting in a incidence rate of 26.1 (95%CI 24.8-27.5).

Overall, the 3 year-probability of treatment failure was of 51.7% (95%CI: 48.5-55.1) for ATV/r, 49.9% (95%CI: 46.6-53.3) for DRV/r and 60.5% (95% CI: 53.2-68.0) for RAL (p=0.158). The 3 year-probability of virological failure 50 was 17.1% (95%CI: 14.4-20.2) for ATV/r, 18.0% (95%CI: 15.3-21.2) for DRV/r and 5.1% (95%: 2.5-10.0) for RAL (p=0.04).

Finally, the 3 year-probability of treatment discontinuation due to toxicity was 21.7% (95%CI: 18.9-24.9) for ATV/r, 13.7% (95%CI: 11.3-16.6) for DRV/r and 4.1% (95% CI: 2.0-8.0) for RAL (p<0.001).

The Kaplan Meier's curves of the risk of experiencing the various end-points, stratified for regimen, are shown in Figure 1.

A total of 627 patients (63.6%) discontinued ATV/r, 605 (59.1%) discontinued DRV/r and 125 (51.9%) RAL. Discontinuation due to toxicity was the main cause of interruption in patients on ATV/r (209 out of 627, 33.3%), while simplification was the main cause of discontinuation both for patients on DRV/r (276 out of 605 discontinuations; 45.6%), and for patients on RAL (59 out of 125 discontinuations, 47.2%) (Table 2).

The main cause of discontinuation were ~~H~~ hyperbilirubinemia for ATV/r, gastrointestinal intolerance and lipid abnormalities for DRV/r. Only 10 patients on RAL discontinued for toxicity, mainly due to allergic reactions, gastrointestinal complaints and nephrotoxicity (Table 2).

<u>Factors associated with the risk of outcomes</u>

After adjusting for age, gender, nation of birth, mode of HIV transmission, hepatitis B and C coinfection, AIDS, baseline CD4 counts and HIV-RNA, year of starting cART and NRTI started, patients given ATV/r showed a 26% statistically significant higher risk of treatment failure (adjusted Hazard Ratio (AHR): 1.26, 95%CI 1.11-1.43 p=0.001) compared to those initiating DRV/r. There was no evidence for a difference in treatment failure among participants starting RAL as compared to those starting DRV/r (AHR 1.02, 95%CI 0.83-1.26 - p=0.83); the risk of treatment failure was lower among patients on RAL as compared to those on ATV/r (AHR 0.81, 95%CI 0.66-0.99 - p=0.05).

Because there was evidence that the proportional hazard assumption might have been violated for this outcome (p=0.06), a sensitivity analysis was performed by including in the model the interaction between the type of treatment and survival time (fitted in the natural logarithmic scale). Results of this analysis were similar, showing again a higher risk of treatment failure in patients starting ATV/r (AHR: 1.26, 95%CI 1.11-1.43 p<0.001) compared to those initiating DRV/r; in contrast, only a trend for lower risk of treatment failure among patients starting RAL as compared to those initiating ATV/r was observed (AHR 0.83, 95%CI 0.67-1.03 - p=0.085).

After controlling for the same set of potential confounding variables, when compared to DRV/r, the probability of virological failure with threshold at 50 copies/ml was significantly lower for people initiating RAL (AHR 0.46, 95%CI 0.24-0.87- p=0.02). The probability of virological failure was also significantly lower for people initiating RAL as compared to those initiating ATV/r (AHR 0.52, 95%CI 0.27-0.99- p=0.05). No differences in virological failure were observed between the two PI/r regimens (ATV/r: AHR 0.85 – 95%CI: 0.66-1.09- vs DRV/r).

Initiation of ATV/r was associated with a higher risk of discontinuation because of toxicity (AHR: 2.09, 95%CI: 1.63-2.67; p<0.001) when compared to DRV/r. Finally, patients who started a RAL-based regimen were less likely to stop due to toxicity as compared to DRV/r (AHR: 0.37, 95%CI: 0.19-0.72; p=0.003) as well as compared to ATV/r (AHR: 0.18, 95%CI: 0.09-0.34; p<0.001) (Table 3).

<u>CD4 count response</u>

A total of 1790 (79.6%) patients had a follow up of at least 2 years, and of these 1747 (97.6%) had ≥1 available CD4 count at 2 year from treatment initiation (808 ATV/r, 796 DRV/r, 143 RAL). Participants reaching 2 years of follow-up and with 2 year-CD4 available were less frequently migrants, HCV and HBV co-infected and more frequently MSM; further, they were less frequently on RAL than patients with a shorter follow up.

Although the three groups started with different median CD4 cell count/cmm (ATV/r 305 DRV/r 254 RAL 369, p<0.001), the mean CD4 recovery was not different among groups (+18.3 [95% -6.0; +42.6] for ATV/r and +10.7 [95%CI -30.7; 52.0] for RAL compared to DRV/r). After adjustment for baseline characteristics, ATV/r showed higher mean CD4 recovery at 2 years (+27.2 [95%CI +2.27; +52.1]) as compared to DRV/r; RAL showed a higher mean CD4 recovery at 2 years as compared to DRV/r, although marginally statistically different (+37.6 [95%CI -3.5; 78.7]).

Discussion

Our analysis substantially confirms and extends to a longer duration of follow-up the results of the ACTG 5257 trial in a clinical setting of HIV-infected persons seen for routine care in Italy.

In detail, our estimates of the incidence of treatment failure according to the three regimens were similar but not identical to those seen in the trial and showed a higher risk of failure for patients starting ATV/r as compared to those initiating the other two regimens. In fact, the absolute estimates of failure in our analysis were considerably higher than those observed in the trial. However, in the trial the definition of treatment failure included virological failure but only discontinuation of drugs due to toxicity/intolerance. We preferred to use a broader definition of treatment failure including the discontinuations of the third drugs for any reasons, given the observational setting of our study and the possible misclassifications of reasons for discontinuation, and this might in part explain the higher frequency of treatment failure in the Icona cohort as compared to that seen in the trial. .

Further, patients from the ICONA cohort were only partially comparable to US patients enrolled in the ACTG trial: in ICONA, there were more subjects who acquired HIV infection by intravenous drug use (8.6% vs 2%) and less subjects who were infected through men to men sexual intercourse (39.7% vs 54%) than in the ACTG trial, reflecting the known differences in the HIV epidemics in Italy vs USA [4]. The different case mix and the real-life setting of the ICONA patients, potentially enriched with a population of less adherent patients, might have also contributed to the higher failure rates seen. .

The probability of discontinuation because of toxicity was higher in our cohort as compared to the ACTG trial, but the trends were similar, with patients who started ATV/r showing the highest risk, DRV/r intermediate risk and RAL the lowest risk. The causes leading to discontinuation because of toxicity of the three drugs are largely expected, with a driving cause represented by hyperbilirubinemia for ATV/r, gastrointestinal complaints for DRV/r and allergic reactions (even if few) for RAL. Also with this respect, our analysis replicates the results seen in the trial.

Further, in our analysis RAL appeared to be superior in terms of tolerability also, although to a less extent, to DRV/r. These data are partly unexpected because patients on RAL showed a higher frequency of comorbidities at treatment initiation. The possible toxic effect of the drug is therefore difficult to disentangle from an apparent channelling bias [7-9]. This was replicated in our multivariable analysis which, after controlling for baseline imbalances between groups, showed identical results.

When we looked at pure virological failure, patients receiving RAL-including combinations showed a 50% reduction in risk of failure as compared to those receiving DRV/r; there was no evidence for a difference in virological failure when comparing the two PI/r against each other. In contrast, the analysis of the trial shows no differences in the rate of virological failure between the three arms regardless of the threshold chosen to define viral failure (50 or 200 copies/mL). Because of the known limitation of adjusting for confounders by multivariable analysis, we cannot rule out that the reduced risk of failure of RAL recipients in our analysis was partly due to this imbalance at baseline.

To our knowledge there are no data verifying the reliability of the ACTG 5257 in clinical settings, even if all regimens have been widely used as first-line. Davis at al [~~14~~10] demonstrated that RAL-based regimens have a lower cost for successfully treated patients compared to DRV/r or ATV/r as first-line regimens in Spain. The STARTMRK [11] demonstrated the high virological potency and tolerability of RAL in naïve patients, with 81% of virologically controlled patients over 96 weeks-follow up. Other information can be derived by observational studies on individual regimens. A recent study from US [12] showed that the probability to be alive and virologically suppressed among patients on RAL was of 71% at 2 years, data not different from what found in our cohort (showing 26% of incidence of treatment failure in a median follow up of 3 years). The Swiss cohort published recently a paper showing few discontinuations due to toxicity in both RAL and dolutegravir-receiving patients [13]. In particular, the main cause of discontinuation for RAL was convenience, similar to our findings showing simplification as main cause of discontinuation. In a previous analysis on late presenters from the ICONA cohort we demonstrated a similar probability of treatment failure in participants on DRV/r and on ATV/r, both resulting in a better response as compared to lopinavir/r given patients [14]. Both DRV/r and ATV/r have been demonstrated to be highly effective in registration trials in comparison to LPV/r [15-16]. In the US setting, there were no differences in the durability of ATV/r and DRV/r regimens [17]. Patients' and physicians' concerns on hyperbilirubinemia together with the availability of other options might have affected the higher probability of treatment failure and discontinuation for toxicity in our data set as compared to previous ones.

Unexpectedly, we found that ATV/r given patients had a better 2-year CD4 recovery as compared to other groups. In contrast, the trial shows a better immune recovery in the RAL arm; there are a number of possible explanations for this discrepancy, including possible selection bias, the relatively small numbers in the RAL group, and, of course, unmeasured confounding.

Our study has several limitations: first, because this is not a randomised study, channelling bias cannot be ruled out; indeed there was an imbalance between treatment arms even in measured potential confounders: for example; RAL was more likely given to participants with less advanced HIV diseases but with more comorbidities. Although we have accounted for these difference in the multivariable analysis, residual confounding might exist.

The major strengths of our analysis are the real life composition of the study population, the possibility to compare the treatment strategy in a setting with free-access to care and the long-term follow-up (on average one year longer than the trial). Indeed, we believe that the most important aspect of our analysis is that it was conducted in Italy so results should be less affected by bias due to socio-economic factors limiting patients' adherence to expensive treatment like in the USA trial setting.

In conclusion, our analysis shows higher absolute risks of failure for all regimens studied compared to those estimated in the randomised comparison but this discrepancy is largely attributable to the difference in the definition of the main endpoint used and the case-mix of the study population. More importantly, the analysis confirms in the real-life setting, the lower tolerability and higher rate of discontinuation of ATV/r compared to DRV/r and RAL observed in the trial. In addition, we found a clear signal that RAL might be superior to both PI/r-based regimens with respect to tolerability and risk of virological failure with a threshold of $\geq$50 copies/mL.

Disclosure Declaration

None of the authors had conflict of interest or receive funding on this work.

Funding

No specific funding for this manuscript. ICONA Foundation is sponsored by unrestricted grants of Bristol Myers Squibb, Gilead, Janssen, Merck Sharp and Dohme and ViiV Healthcare Italy.

References

1. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Department of Health and Human Services. (Last updated July 14, 2016). Available at https://aidsinfo.nih.gov/contentfiles/lvguidelines/adultandadolescentgl.pdf
2. Günthard HF, Saag MS, Benson CA, et al. Antiretroviral Drugs for Treatment and Prevention of HIV Infection in Adults. *JAMA* 2016;**316**(2):191.
3. EACS Guidelines for clinical management and treatment of HIV infected adults and adolescents. Version 8.1. Available: www.eacssociety.org.
4. Lennox JL, Landovitz RJ, Ribaudo HJ, et al. Efficacy and Tolerability of 3 Nonnucleoside Reverse Transcriptase Inhibitor–Sparing Antiretroviral Regimens for Treatment-Naive Volunteers Infected With HIV-1. *Ann Intern Med* 2014; **161**(7):461.
5. d'Arminio Monforte A, Lepri AC, Rezza G, et al. Insights into the reasons for discontinuation of the first highly active antiretroviral therapy (HAART) regimen in a cohort of antiretroviral naïve patients. I.CO.N.A. Study Group. Italian Cohort of Antiretroviral-Naïve Patients. *AIDS* 2000; **14**(5):499-507.
6. d'Arminio Monforte A, Cozzi-Lepri A, Girardi E, et al. Late presenters in new HIV diagnoses from an Italian cohort of HIV-infected patients: prevalence and clinical outcome. *Antivir Ther* 2011; **16**(7):1103-1112.
7. Bañón S, Machuca I, Araujo S, et al. Efficacy, safety, and lack of interactions with the use of raltegravir in HIV-infected patients undergoing antineoplastic chemotherapy. *J Int AIDS Soc* 2014; **17**(4 Suppl 3):19590 .

8. Lennox JL, DeJesus E, Lazzarin A, et al. Safety and efficacy of raltegravir-based versus efavirenz-based combination therapy in treatment-naive patients with HIV-1 infection: a multicentre, double-blind randomised controlled trial. *Lancet* 2009; **374**(9692):796-806.
9. Pandey KK. Raltegravir in HIV-1 Infection: Safety and Efficacy in Treatment-Naïve Patients. Clin Med Rev Ther 2011; 2012(4):13–30.
10. Davis AE, Brogan AJ, Goodwin B, Nocea G, Lozaro V: Short-term cost and efficiency analysis of raltegravir versus atazanavir/ritonavir or darunavir/ritonavir for treatment-naïve adults with HIV-1 infection in Spain. HIV Clin Trials 2017; 18 (5-6): 214-22
11. Rockstroh JK, DeJesus E, Lennox JL, et al; STARTMRK Investigators. Durable efficacy and safety of raltegravir versus efavirenz when combined with tenofovir/emtricitabine in treatment-naive HIV-1-infected patients: final 5-year results from STARTMRK. *J Acquir Immune Defic Syndr*. 2013;63(1):77-85.
12. Edwards JK, Cole SR, Hall HI, Mathews WC, Moore RD, Mugavero MJ, Eron JJ;CNCIS investigators. Virologic suppression and CD4 cell count recovery after initiation of raltegravir- or efavirenz- containing HIV treatment regimens. *AIDS* 2018;32: 261-266
13. Elzi L, Erb S, Furrer H, et al for the Swiss HIV Cohort Study Group Adverse events of raltegravir and dolutegravir. *AIDS* 2017; 31 (13): 1853–185
14. d'Arminio Monforte A, Cozzi-Lepri A, Maggiolo F, et al for the IconaFoundation Study cohort. Response to First-Line Ritonavir-Boosted Protease Inhibitors (PI/r)-Based Regimens in HIV Positive Patients Presenting to Care with Low CD4 Counts: Data from the Icona Foundation Cohort. *PLoS ONE* 2016; 11(6): e0156360. doi:10.1371/journal
15. Molina JM, Andrade-Villanueva J, Echevarria J, et al CASTLE Study Team: Once-daily atazanavir/ritonavir versus twice-daily lopinavir/ritonavir, each in combination with tenofovir and emtricitabine, for management of antiretroviral-naive HIV-1-infected patients: 48 week efficacy and safety results of the CASTLE study. *Lancet.* 2008; 372(9639):646-55.

16. Ortiz R, DeJesus E, Khanlou H, et al: ARTEMIS Study: Efficacy and safety of once-daily darunavir/ritonavir versus lopinavir/ritonavir in treatment-naive HIV-1-infected patients at week 48. *AIDS* 2008; 22(12): 1389-1397

17. Farr AM, Johnston SS, Ritchings C, Brouillette M, Rosenblatt L. No difference in persistence to treatment with atazanavir or darunavir in HIV patients in a real-world setting. *J Int AIDS Soc.* 2014 Nov 2;17(4 Suppl 3):19538.

Figure 1. Kaplan Meier curves estimating cumulative probability of various end-points according to drug regimens started.

Figure

Click here to download Figure Fig1_HCT_rev2 def.tif

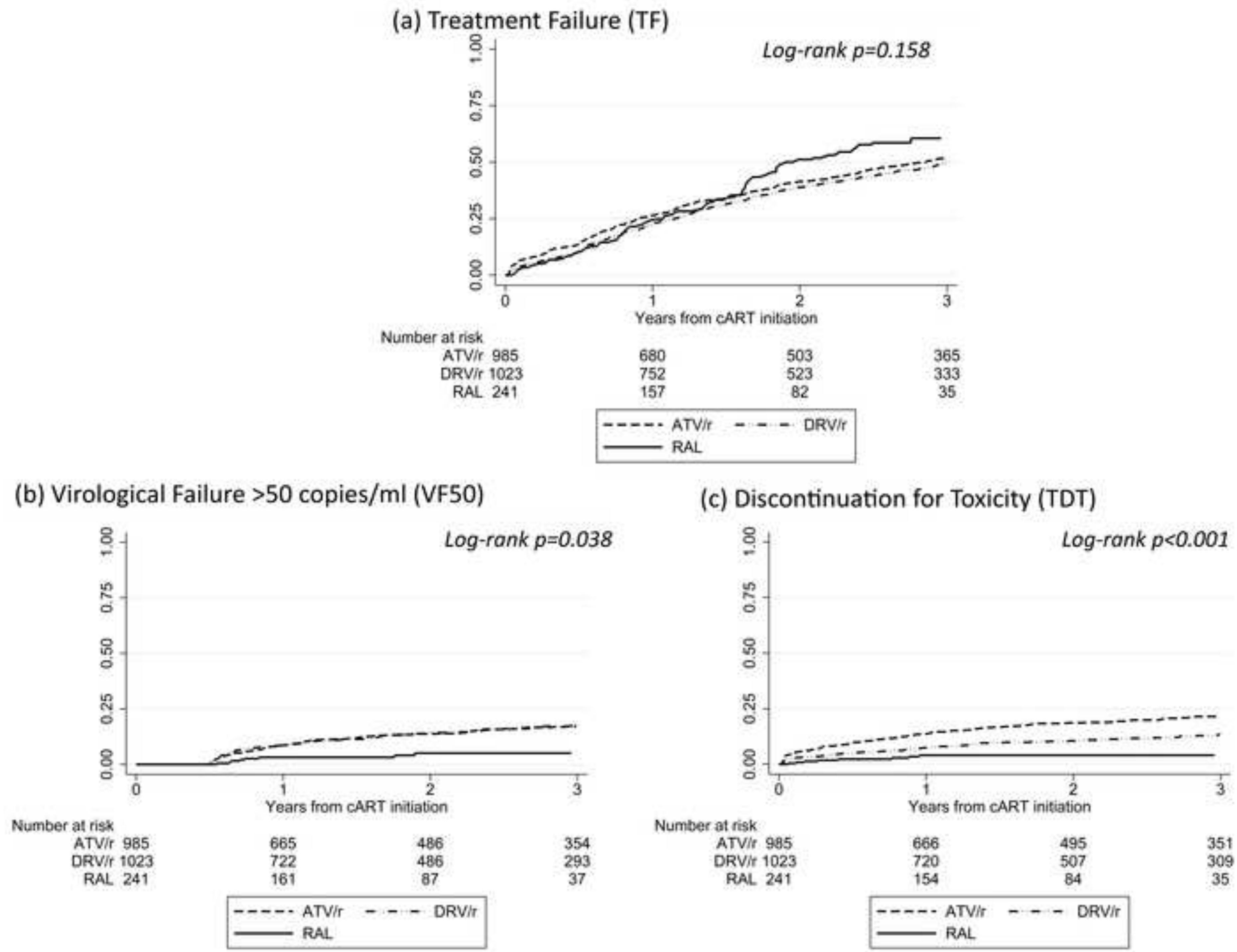

table 1 no tracks

Table 1. Main characteristics of 2,249 patients according to the third drug started at their first antiretroviral regimen

	ATV/r N=985	DRV/r N=1,023	RAL N=241	p-value*	Total N=2,249
Gender, n(%)					
Male	745 (75.6%)	835 (81.6%)	196 (81.3%)	0.003	1,776 (79.0%)
Age, yrs, median (IQR)	39 (32-47)	40 (33-49)	43 (35-50)	<0.001	40 (32-48)
Migrants, n (%)	240 (24.4%)	209 (20.4%)	42 (17.4%)	0.022	491 (21.8%)
Mode of HIV transmission, n(%)					
Heterosexual	450 (45.7%)	426 (41.6%)	106 (44.0%)	<0.001	982 (43.7%)
IDU	118 (12.0%)	62 (6.1%)	14 (5.8%)		194 (8.6%)
MSM	354 (35.9%)	436 (42.6%)	102 (42.3%)		892 (39.7%)
Other/unknown	63 (6.4%)	99 (9.7%)	19 (7.9%)		181 (8.0%)
AIDS diagnosis, n(%)	88 (8.9%)	164 (16.0%)	29 (12.0%)	<0.001	281 (12.5%)
≥1 Comorbidity, n(%)	42 (4.3%)	52 (5.1%)	24 (10.0%)	0.002	118 (5.2%)
Time from HIV diagnosis to first cART, months, median (IQR)	4 (1-32)	2 (1-17)	3 (1-24)	<0.001	3 (1-24)
HCV co-infection, n(%)					
Positive	125 (12.7%)	80 (7.8%)	19 (7.9%)	0.001	224 (10.0%)
Negative	769 (78.1%)	830 (81.1%)	188 (78.0%)		1787 (79.5%)
Not tested	91 (9.2%)	113 (11.1%)	34 (14.1%)		238 (10.5%)
HBV co-infection, n(%)					
Positive	41 (4.2%)	37 (3.6%)	14 (5.8%)	0.311	92 (4.1%)
Negative	818 (83.1%)	833 (81.4%)	190 (78.8%)		1841 (81.9%)
Not tested	126 (12.8%)	153 (15.0%)	37 (15.4%)		316 (14.0%)
CD4 cell/cmm, n (%)					
0-200	312 (31.7%)	443 (43.3%)	68 (28.2%)	<0.001	823 (36.6%)
201-350	299 (30.4%)	228 (22.3%)	49 (20.3%)		576 (25.6%)
351-500	218 (22.1%)	207 (20.2%)	48 (19.9%)		473 (21.0%)
>501	135 (13.7%)	120 (11.7%)	66 (27.4%)		321 (14.3%)
Not available	21 (2.1%)	25 (2.4%)	10 (4.2%)		56 (2.5%)
CD4 cell/cmm, mean (SD)	306 (205)	263 (210)	375 (273)	<0.001	294 (218)
CD4 cell/cmm, median (IQR)	300 (152-410)	244 (80-394)	346 (153-532)	<0.001	277 (120-415)
HIV RNA copies/mL, n(%)					
50-20,000	247 (25.1%)	210 (20.5%)	68 (28.2%)	0.001	525 (23.3%)
20,000-100,000	308 (31.3%)	269 (26.3%)	78 (32.4%)		655 (29.1%)
100,000-250,000	181 (18.4%)	209 (20.4%)	34 (14.1%)		424 (18.8%)
>250,000	213 (21.6%)	301 (29.4%)	51 (21.2%)		565 (25.1%)
Not available	36 (3.7%)	34 (3.3%)	10 (4.2%)		80 (3.6%)
HIV RNA log_{10} copies/mL, median (IQR)	4.8 (4.3-5.3)	5.0 (4.5-5.5)	4.8 (4.2-5.3)	<0.001	4.9 (4.3-5.4)
Calendar year of cART start, n(%)					
2008-2009	98 (9.9%)	12 (1.2%)	14 (5.8%)	<0.001	124 (5.5%)
2010-2011	354 (35.9%)	265 (28.7%)	28 (11.6%)		647 (28.8%)
2012-2013	356 (36.1%)	403 (39.4%)	52 (21.6%)		811 (36.1%)
2014-2015	177 (18.0%)	343 (33.5%)	147 (61.0%)		667 (29.7%)
NRTI pair, n(%)					
Tenofovir/Emtricitabine	852 (86.5%)	886 (86.6%)	207 (85.9%)	0.958	1945 (86.5%)
Abacavir/Lamivudine	133 (13.5%)	137 (13.4%)	34 (14.1%)		304 3.5%)

* Chi-square or Kruskal-Wallis test as appropriate

NRTI=nucleoside reverse transcriptase inhibitors

IDU=intravenous drug addicts

MSM=men sex with men

Table 2. All causes of discontinuation and details of causes of discontinuation due to toxicity according to the regimen given

All Causes of Discontinuation	**ATV/r N=627**	**DRV/r N=605**	**RAL N=125**	**Total N=1357**
Simplification	184 (29.4%)	276 (45.6%)	59 (47.2%)	519 (38.2%)
Toxicity	209 (33.3%)	124 (20.5%)	10 (8.0%)	343 (25.3%)
Other	70 (11.2%)	72 (11.9%)	11 (8.8%)	153 (11.3%)
Missing	38 (6.1%)	39 (6.5%)	9 (7.2%)	86 (6.3%)
Failure	50 (8.0%)	26 (4.3%)	7 (5.6%)	83 (6.1%)
Patient's decision	39 (6.2%)	23 (3.8%)	11 (8.8%)	73 (5.4%)
Clinical trial	14 (2.2%)	26 (4.3%)	11 (8.8%)	51 (3.8%)
Structured Treatment Interruption	18 (2.9%)	13 (2.2%)	6 (4.8%)	37 (2.7%)
Pregnancy	4 (0.6%)	4 (0.7%)	1 (0.8%)	9 (0.7%)
Death	1 (0.2%)	2 (0.3%)	0 (0.0%)	3 (0.2%)

Causes of Discontinuation due to Toxicity	**ATV/r N=209**	**DRV/r N=124**	**RAL N=10**	**Total N=343**
Gastrointestinal Toxicity	31 (14.8%)	35 (28.2%)	2 (20.0%)	68 (19.8%)
Hyperbilirubinemia	58 (27.8%)	0 (0.0%)	0 (0.0%)	58 (16.9%)
Allergic Reactions / Rash	26 (12.4%)	24 (19.3%)	2 (20.0%)	52 (15.2%)
Lipid Metabolism Toxicity	15 (7.2%)	35 (28.2%)	0 (0.0%)	50 (14.6%)
Others	20 (9.6%)	15 (12.1%)	3 (30.0%)	38 (11.1%)
Hepatotoxicity *	28 (13.4%)	6 (4.8%)	0 (0.0%)	34 (9.9%)
Nephroxicity	23 (11.0%)	6 (4.8%)	2 (20.0%)	31 (9.0%)
Osteopenia / Osteoporosis	4 (1.9%)	3 (2.4%)	1 (10.0%)	8 (2.3%)
Toxicity Not Specified	4 (1.9%)	0 (0.0%)	0 (0.0%)	4 (1.2%)

**Hepatotoxicity other than hyperbilirubinemia* table 3 no tracks

Table 3. Hazard ratio from fitting three separate Cox regression models.

	# event	PYFU	Crude HR (95%CI)	p-value	Adjusted* HR (95%CI)	p-value
TF (HIV-RNA>200 copies/mL or discontinuation)						
DRV/r	623 (43 VF200, 580 D)	2504	1.00		1.00	
ATV/r	679 (65 VF200, 614 D)	2497	1.08 (0.96-1.22)	0.200	1.26 (1.11-1.43)	0.001
RAL	131 (3 VF200, 128 D)	430	1.17 (0.96-1.42)	0.129	1.02 (0.83-1.26)	0.833
VF50 (HIV-RNA>50 copies/mL)						
DRV/r	149	2325	1.00		1.00	
ATV/r	154	2426	0.85 (0.66-1.09)	0.212	0.88 (0.67-1.15)	0.345
RAL	11	440	0.38 (0.20-0.71)	0.003	0.46 (0.24-0.87)	0.018
Discontinuation due to toxicity						
DRV/r	124	2351	1.00		1.00	
ATV/r	209	2403	1.79 (1.42-2.27)	<0.001	2.09 (1.63-2.67)	<0.001
RAL	10	422	0.42 (0.22-0.81)	0.010	0.37 (0.19-0.72)	0.003

*Each model adjusted for age, gender, nation of birth, mode of HIV transmission, hepatitis co-infection status, AIDS diagnosis, nucleoside pair started, baseline CD4 count and viral load and year of starting cART.

(TF= treatment failure, VF=virological failure, VF200=HIV-RNA>200 copies/mL, D=discontinuation, PYFU=person-years follow-up, HR=hazard ratio).