Population dynamics of HIV drug resistance among pre-treatment and treatment-experienced persons with HIV during treatment scale-up in Uganda: a population-based longitudinal study

Michael A. Martin[1], Steven James Reynolds[2,3,4], Charles Ssuuna[2], Brian T. Foley[5], Fred Nalugoda[2], Thomas C. Quinn[2,3,4], Steven A. Kemp[6], Margaret Nakalanzi[2], Edward Nelson Kankaka[2], Godfrey Kigozi[2], Robert Ssekubugu[2], Ravindra K. Gupta[6,7], Lucie Abeler-Dörner[8], Joseph Kagaayi[2,9], Oliver Ratmann[10], Christophe Fraser[8,11], Ronald Moses Galiwango[2], David Bonsall[8,11,*], M. Kate Grabowski[1,2,13,*], on behalf of the PANGEA-HIV Consortium and the Rakai Health Sciences Program

[1]Department of Pathology, Johns Hopkins School of Medicine, Baltimore, MD, USA
[2]Rakai Health Sciences Program, Kalisizo, Uganda
[3]Division of Infectious Disease, Department of Medicine, Johns Hopkins School of Medicine, Baltimore, MD, USA
[4]Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA
[5]Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, New Mexico, USA
[6]Department of Medicine, University of Cambridge, Cambridge, UK
[7]Africa Health Research Institute, KwaZulu-Natal, South Africa
[8]Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Medicine, University of Oxford, Oxford, UK
[9]Makerere University School of Public Health, Kampala, Uganda
[10]Department of Mathematics, Imperial College London, London, England, United Kingdom
[11]Pandemic Sciences Institute, University of Oxford, Oxford, UK
[12]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK
[13]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

[*]These authors contributed equally

[*]Corresponding author: mgrabow2@jhu.edu

Abstract

Background

Longitudinal data on the population prevalence of HIV drug resistance during scale-up of HIV treatment in Africa are extremely limited. We estimated trends in HIV drug resistance prevalence during ART program expansion from a population-based surveillance cohort in southern Uganda.

Methods

We analyzed data from Rakai Community Cohort Study participants aged 15-49 during four survey rounds conducted between 2012 (round 15) and 2019 (round 19). Consenting participants were tested for HIV and completed questionnaires. Persons living with HIV (PLHIV) provided samples for viral load quantification and virus deep-sequencing. Sequence data were used to predict resistance profiles. The prevalence of class-specific resistance and resistance-conferring substitutions were estimated using robust log-Poisson regression.

Findings

93,659 participant visits were contributed between 2012 and 2019, including 17,471 (18.65%) from PLHIV. Using deep-sequencing data from 3,713 pre-treatment participant-visits we estimated that the population prevalence of viremic NNRTI, NRTI, and PI resistance decreased significantly between 2012 and 2017 (PR = 0.38, 95% CI 0.25 – 0.57; 0.20, 95% CI 0.09 – 0.45; 0.19, 95% CI 0.09 – 0.39, respectively) with increasing viral suppression. Among viremic pre-treatment PLHIV, the prevalence of NNRTI resistance increased two-fold (PR = 1.96, 95% CI 1.31-2.95) to 9.77% (7.35% - 12.97%) over the same time period. We did not observe an increase in NRTI or PI resistance in this population. The 2017 prevalence of NNRTI and NRTI resistance among viremic treatment-experienced PLHIV was 47.67% (95% CI 40.94% - 55.50%) and 36.55% (95% CI 30.14% - 44.31%), respectively. Single-class resistance predominated among resistant pre-treatment PLHIV (83.05%) whereas most treatment-experienced resistance was multi-class (76.65%). In 2017, 10.13% (95% CI 7.83%-13.63%) and 9.98% (95% CI 6.43%-15.51%) of viremic pre-treatment and treatment-experienced PLHIV harbored the inT97A mutation.

Interpretation

Prevalence of HIV drug resistance among viremic PLHIV significantly increased with scale-up of ART programs. The prevalence of inT97A is potentially concerning considering the recent roll-out of dolutegravir-based regimens.

Funding

National Institutes of Health, the Bill & Melinda Gates Foundation, and the U.S. President's Emergence Plan for AIDS Relief through the Centers for Disease Control and Prevention.

Research in context

Evidence before the study

We searched PubMed for studies matching the keywords "hiv" "resistance" "longitudinal" "cohort" "population" published since 2004 (the beginning of antiretroviral therapy (ART) availability in sub-Saharan Africa) and identified 48 studies. We excluded 33 studies not based in sub-Saharan Africa, four studies primarily concerned with coinfection with other pathogens (e.g. HBV, *M. tuberculosis*), two studies concerned with insulin resistance, one sequencing-methods paper, and one paper concerned with host susceptibility to HIV infection. The remaining seven studies were not population-based meaning that the study population was not all persons but e.g. people living with HIV enrolled in care at a given clinic. We identified no previous longitudinal population-based cohort studies of HIV ART resistance in sub-Saharan Africa.

Added value of this study

We estimated the prevalence of drug resistance over four survey rounds of a population-based open-cohort study in southern Uganda between 2012 and 2019 during a period of intense treatment scale-up. We show that pre-treatment resistance to non-nucleoside reverse transcriptase inhibitors (NNRTIs) among pre-treatment PLHIV increased significantly during the scale-up of ART. We further show that among viremic treatment-experienced individuals 48% and 37% harbored resistance to NNRTIs and nucleoside-reverse transcriptase inhibitors (NRTIs), the majority of which harbored multiclass resistance. While drug resistance among people living with viremic HIV increased, the overall prevalence of viremic HIV drug resistance in the population decreased by about two-thirds due to increasing population viral load suppression. The most common resistance mutation in our population was inT97A, a known compensatory mutation for integrase strand transfer inhibitor (INSTI) resistance. In contrast to other mutations, presence of inT97A did not depend on treatment status.

These results provide the first longitudinal population-based estimates of temporal trends in the prevalence of drug resistance during ART program expansion in a high-burden setting. Further, they provide critical insight into the landscape of prevalent drug resistance substitutions circulating in this population.

Implications of all the available evidence

Scale-up of HIV treatment has increased the prevalence of drug resistance mutations among viremic people living with HIV in sub-Saharan Africa. The relatively high prevalence of NNRTI resistance has prompted a recent shift to first-line regimens including dolutegravir (an INSTI) in combination with NRTIs. The high prevalence of an INSTI compensatory mutation in our population further warrants continuing monitoring of treatment failures and the prevalence of drug resistance in high burden settings.

Introduction

Antiretroviral therapy (ART) effectively suppresses human immunodeficiency virus (HIV) replication in persons living with HIV (PLHIV),[1] which slows disease progression[2] and prevents viral transmission.[3] With increased uptake of ART as well as interventions such as voluntary medical male circumcision, global annual HIV incidence has fallen by nearly 40% since 2010.[4]

Viral resistance to ART can offset the clinical and public health impact of treatment scale-up. Drug resistance can be acquired when an individual initially infected with a susceptible viral population develops resistance following treatment exposure. This is more common when treatment adherence is intermittent,[5] but can occur despite high adherence.[6] In low- and middle-income countries approximately 15% of individuals initiating therapy fail to achieve suppression,[7] which is largely due to resistance.[8] Viral genomes with resistance-conferring mutations can be transmitted to individuals without a history of treatment, increasing their risk of failing first-line treatment approximately five-fold.[9]

Cross-sectional studies in sub-Saharan Africa have revealed an increase in the prevalence of transmitted non-nucleoside reverse transcriptase inhibitor (NNRTI) resistance during the roll-out of ART. A systematic meta-regression estimated that pre-treatment NNRTI resistance increased by 1.3% annually between 2001 and 2016 in this setting, reaching 10.1% in eastern Africa by 2016.[10] Similarly, a systematic meta-analysis of studies conducted between 2017 and 2022 in eastern Africa estimated the pooled prevalence of pre-treatment NNRTI resistance to be 9.4%.[11] Longitudinal surveys of PLHIV who report to PEPFAR-supported clinics in Uganda, Kenya, Tanzania, and Nigeria (African Cohort Study) have further shown an increase in pre-treatment resistance from 8.8% (2013-2015) to 16.2% (2016-2019), of which most was NNRTI resistance.[12] However, aggregation of findings across populations with different access to diagnostics and treatment and different clinical guidelines, limited temporal resolution, and a focus on PLHIV who are enrolled in care limit the generalizability of prior studies and introduce uncertainty in estimates of the burden of HIV drug resistance among PLHIV.

Overall, there is a critical lack of longitudinal estimates of the prevalence of HIV drug resistance in sub-Saharan Africa from population-based studies spanning the scale-up of ART. These studies are essential to understanding the interaction between increased treatment uptake and viral evolution and the implications of this interaction for the clinical and population burden of HIV.

Here, we used HIV deep-sequence data from 4,130 participant-visits collected over a nine-year period from 3,456 PLHIV spanning a period of intense ART scale-up to identify temporal trends in HIV drug resistance among pre-treatment and treatment experienced PLHIV in a population-based open cohort study in southern Uganda. Through the study, first-line NNRTI-based regimens were the standard of care, often in combination with two NRTIs.[13] We estimated the prevalence of NNRTI, nucleoside reverse transcriptase inhibitor (NRTI), and protease inhibitor (PI) resistance among PLHIV and among the entire study population using data from 3,713 pre-treatment and 417 treatment-experienced PLHIV. We further evaluated temporal trends in transmitted NNRTI resistance among 223 HIV-1 seroconverters. Finally, we assessed the temporal dynamics of viral substitutions associated with drug resistance in pre-treatment and treatment-experienced PLHIV.

Methods

Study design and participant selection

The RCCS is an open population-based census and cohort study conducted at approximately 18-24 month intervals (appendix pp 2-3) in agrarian (HIV prevalence 9-26%[14]), semi-urban trading (11-21%[14]), and Lake Victoria fishing (38-43%[14]) communities in southern Uganda.[15] At survey rounds households in participating communities are censused and residents aged 15-49 capable of providing informed consent (or assent if under 18) are invited to participate. Consenting participants are administered a structured questionnaire that obtains sociodemographic, behavioral, and health information, including self-reported past and current ART use. Voluntary HIV testing of participants is conducted using a rapid test algorithm[16] and venous blood samples are taken for viral quantification and sequencing. Participants with serologically confirmed HIV infection were considered pre-treatment during a given round if they reported never having taken ARTs at that round and all prior rounds in which they participated. PLHIV were considered treatment-experienced during a given round if they reported using ART at that round or any earlier rounds.

Here, we used RCCS survey data from survey rounds 1 through 19, conducted between November 5, 1994 and November 4, 2020, and HIV viral load and sequence data from survey round 15, which began on August 10, 2011, through round 19 collected from a total of 43 communities. As they occur over a multi-year window, study rounds were plotted at the median date over which the round occurred and referred to by the year of the median date (appendix p 2). Reporting of this study adheres to the STROBE guidance[17].

HIV viral load quantification

HIV viral load was measured on serum/plasma samples using the Abbott real-time m2000 assay (Abbott Laboratories). Viral load measurements were available for 1,959/3,499 (55.99%) of PLHIV in round 15 and 13,979/14,025 (99.67%) PLHIV in rounds 16 through 19. Participant visits from PLHIV in rounds 16 through 19 without viral load measurements (46/14,025, 0.33%) were dropped. Viral loads ≥ 1000 copies/mL were considered viremic. As 1,783/2,380 of pre-treatment participant-visits in round 16 through 19 were viremic, pre-treatment PLHIV in round 15 with missing viral load measurements were assumed to be viremic.

HIV deep sequencing

Full-length HIV deep sequencing was conducted through the Phylogenetics and Networks for Generalized HIV Epidemics in Africa consortium (PANGEA-HIV).[18,19] As described elsewhere,[20] for survey rounds 15 and 16 sequencing on Illumina MiSeq and HiSeq platforms was attempted for participants who self-reported never having been on ART and either had a missing viral load or were known to be viremic (appendix p 4). For survey rounds 17 and 18 and select round 15 and 16 participant-visits, a total of 46.92% of all participant-visits (appendix p 7), sequencing of viremic samples was attempted using the veSEQ-HIV protocol.[21] For round 19 sequencing data was only available for samples collected through May 17, 2022 (N = 170/453, 37.53% of participant visits) and therefore round 18 was used as an end-point for longitudinal deep-sequence based analyses. Consensus sequences were generated using shiver[22] and subtyped by identifying the most similar reference sequence and using the Recombination Identification Program.[23]

Identification of drug resistance substitutions

A validated bioinformatic pipeline, drmSEQ, was used to identify amino acid substitutions associated with reduced susceptibility to ART and predict individual drug class susceptibilities.[24] Paired-end reads were trimmed of adapters, primers, and low-quality bases and then filtered to remove *pol* hypermutated sequences and non-HIV *pol* sequences. Reads were deduplicated using Picard MarkDuplicates[25] and locally aligned to 142 HIV subtype references using blastx.[26] A manually-curated codon-restricted multiple-alignment of the references was used to lookup coordinates and mutations relative to HXB2 (GenBank: K03455.1). Only mutations supported by a minimum of 10 PCR-deduplicated reads and by ≥5% of reads spanning the corresponding site were considered.[27] Amino acid substitutions were scored according to the Stanford University HIV Drug Resistance Database. Scores were summed to predict susceptibility to 25 HIV drugs (appendix p 8).[28–30] A score ≥30 (intermediate/high-level) for a given drug was categorized as resistant. Resistance was not predicted if less than half of the relevant positions for a given drug had fewer than 10 reads. Samples with resistance to at least one drug within each class were categorized as resistant to that class. Samples in which there was insufficient sequencing coverage for one or more drug within a class were not assigned a resistance categorization for that class.

Outcome measures

The primary outcomes of this study were the prevalence of NNRTI, NRTI, and PI resistant genotypes among viremic pre-treatment and treatment-experienced PLHIV in the RCCS stratified by survey round. The overall RCCS population prevalence of viremic resistance, stratified by survey round and self-reported treatment status was also reported. We further estimated the prevalence of multi-class resistance among viremic round 18 participant visits with sufficient sequence data to predict resistance to all NNRTIs, NRTI, and PIs. The prevalence of resistance-conferring viral substitutions was estimated in round 15 through 19 among pre-treatment PLHIV with resistance predictions to all drugs and in rounds 17 and 18 among treatment-experienced PLHIV with resistance predictions to all drugs. For context, we estimated the population prevalence of PLHIV, viremic PLHIV, viremic PLHIV pre-treatment, and viremic treatment-experienced PLHIV in each round.

A sub-analysis was conducted among seroincident cases who were HIV negative at one RCCS visit and positive in the next during rounds 15 through 19 and self-reported never having been exposed to ART at their first positive RCCS visit. Among seroincident individuals with available genotypic NNRTI resistance data, we estimated the probability of harboring NNRTI resistance as a function of seroconversion round.

We used inverse probability weighting based on availability of a viral load measurement (True/False), log10 copies/mL among samples with viral load measurements, community type (agrarian/fishing/trading, except for in the seroincident analysis due to small sample size), age category ((14,24]/(24,34]/(34,50]), and sex (M/F) stratified by survey round to account for missing sequence data among some viremic participant-visits and insufficient coverage to predict resistance among some sequenced participant-visits. See appendix pp 4-16 for detailed methods.

Statistical methods

Statistical analyses were conducted in R v.4.3.1.[31] Prevalence was estimated using Poisson regression with a log-link and robust (sandwich) standard errors[32] which were fit using the GLM function in the native stats package, sandwich v.3.0.2[33] and lmtest v.0.9.40.[34] 95% confidence intervals and p-values ($\alpha = 0.05$) were calculated using the Wald method. χ^2 p-values ($\alpha = 0.05$) were calculated using the stats package in R. Repeated measures were evaluated using geepack v.1.3.9[35] and an independent correlation structure was assumed as it generated estimates with the widest confidence intervals (appendix pp 17-21). Data analysis and visualization was done using tidyverse v.2.0.0,[36] ggplo2 v.3.4.3,[37] cowplot v.1.1.1,[38] patchwork v. 1.1.3[39], ggpattern v.1.0.1,[40] and ggrepel v.0.9.3.[41] Readxl v.1.4.3[42] and haven v.2.5.3[43] were used to parse data files.

Ethics

The RCCS is administered by the Rakai Health Sciences Program (RHSP) and has received ethical approval from the Uganda Virus Research Institute's Research and Ethics Committee (HS540), the Uganda National Council for Science and Technology (GC/127/15/11/137), and the Western Institutional Review Board, Olympia, WA (20031318). Participants provided written informed consent at each survey round.

Role of the funding source

The funders had no role in study design, collection, analysis, and interpretation of data; and no role in the writing of the report and decision to publish.

Results

Study population

Between August 10, 2011 and November 4, 2020, a total of 43,369 people participated in the RCCS, of whom 7,925 (18.27%) were PLHIV. Overall, the data contains 93,659 participant-visits, of which about a fifth were from PLHIV (table 1). Among participant-visits from PLHIV, 34.07% were contributed by viremic PLHIV and of these 86.96% were pre-treatment. A total of 320 PLHIV were identified as seroconverting with a last negative and first positive HIV test in consecutive survey rounds.

Overall, HIV seroprevalence among RCCS participants slightly decreased from 20.37% (95% CI 19.78% - 20.98%) in round 15 to 17.20% (95% CI 16.68% - 17.74%) in round 19 (figure 1A, appendix p 22), while population prevalence of HIV viremia among persons with and without HIV decreased considerably from 8.14% (95% CI 7.75% - 8.55%) to 2.34% (95% CI 2.14% - 2.57%) between round 16 and 19. These declines were accompanied by a nearly nine-fold decrease in the population prevalence of pre-treatment viremia over the study. The population prevalence of treatment-experienced viremia remained stable at around 1%.

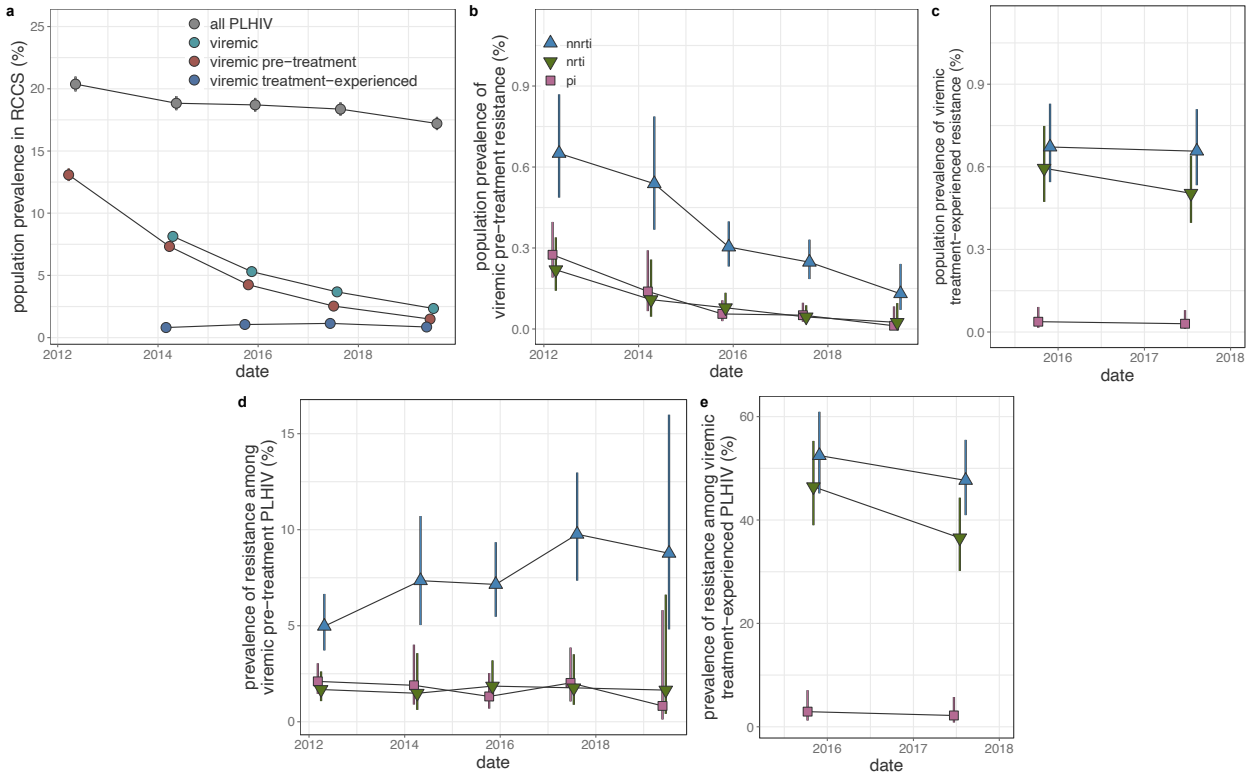


Figure 1: Longitudinal trends in HIV seroprevalence, population prevalence of pretreated HIV, and prevalence of drug resistant HIV in the Rakai Community Cohort Study, 2012-2019. (A) Estimated prevalence of all HIV, viremic HIV, viremic pre-treatment HIV, and viremic treatment-experienced HIV in each round. Due to missing viral load data, prevalence of viremic HIV and viremic treatment-experienced HIV are not estimated in the 2012 survey. (B, C) Estimated overall population prevalence of pre-treatment (B) and treatment-experienced (C) NNRTI, NRTI, and PI resistance. (D, E) Estimated prevalence of NNRTI, NRTI, and PI resistance among viremic pre-treatment PLHIV (D) and treatment-experienced (E) PLHIV. Estimates were generated using robust log-Poisson regression with robust standard errors. Error bars indicate the Wald 95% confidence interval for the mean value within each category. For visual clarity, points are jittered along the x-axis. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = integrase inhibitors.

Identification of resistance genotypes in deep-sequence data

Sufficient sequence data were available to reliably genotype 4,130/5,953 (69.38%) viremic participant visits (3,713/5,177 [71.72%] pre-treatment and 417/776 [53.74%] treatment-experienced, table 2) for at least one drug. A total of 43.56%, 51.74%, 51.10%, and 59.92% were successfully genotyped for all INSTIs, NNRTIs, NRTIs, and PIs, respectively (appendix pp 23-26). Of those, <1% (INSTI), 12.37% (NNRTI), 6.71% (NRTI), and 1.85% (PI) had resistant genotypes (appendix p 27). Given the low number of viremic samples with INSTI resistant genotypes and the limited INSTI use in this setting over the study period we did not estimate the prevalence of INSTI resistance in downstream analyses. Genotyping was successful for all INSTIs, NNRTIs, NRTIs, and PIs for 39.24% of viremic participant visits (appendix p 29). Missing genotype data was more common among samples with missing viral loads or with lower viral loads when available and was slightly more common for female participants.

Population prevalence of resistant genotypes

205 During round 15, the overall population prevalence of viremic pre-treatment NNRTI, NRTI, and
206 PI resistant genotypes was estimated to be 0.65% (95% CI 0.49% - 0.87%), 0.22% (95% CI
207 0.14% - 0.34%), and 0.27% (95% CI 0.19% - 0.40%), respectively (figure 1B, appendix p 30).
208 Due to the decrease in the population prevalence of pre-treatment viremia, the overall population
209 prevalence of NNRTI, NRTI, and PI resistant genotypes decreased significantly between study
210 round 15 and 18 to 0.25% (95% CI 0.19% - 0.33%), 0.04% (95% CI 0.02% - 0.09%), and 0.05%
211 (95% CI 0.03% - 0.10%), respectively.
212
213 In survey rounds 17 and 18 where sequence data were available for viremic pre-treatment and
214 treatment-experienced PLHIV, the majority of participant visits with resistant genotypes were
215 contributed by treatment-experienced PLHIV (appendix p 31). Specifically in round 18, 0.66%
216 (95% CI 0.53%-0.81%) and 0.50% (95% CI 0.40%-0.64%) of participant visits were contributed
217 by viremic treatment-experienced PLHIV with NNRTI and NRTI resistant genotypes,
218 respectively (figure 1C, appendix p 32). Compared to PLHIV pre-treatment, treatment-
219 experienced PLHIV contributed 2.64- and 12.50-fold more NNRTI and NRTI resistant
220 genotypes. The total round 18 population prevalence of NNRTI and NRTI resistant genotypes,
221 regardless of treatment status, was 0.90% (95% CI 0.76%-1.06%) and 0.51% (95% CI 0.41%-
222 0.64%), respectively (appendix p 33). The overall population prevalence of resistant genotypes
223 was stable between round 17 and 18.
224
225 *Prevalence of resistant genotypes among PLHIV*
226 Despite the observed decrease in the population prevalence of resistant genotypes among pre-
227 treatment PLHIV, this decrease was slower than that of all pre-treatment viremia. Consequently,
228 the prevalence of NNRTI resistant genotypes among pre-treatment PLHIV approximately
229 doubled between round 15 and round 18 (prevalence ratio 1.96, 95% CI 1.31-2.95, figure 1C,
230 table 3), reaching an estimated prevalence of 9.77% (95% CI 7.35% - 12.97%). This increase
231 was observed among both men and women, across all age groups, and within all community
232 types (appendix pp 34-36). In contrast, the prevalence of NRTI and PI resistant genotypes among
233 viremic pre-treatment PLHIV remained stable and below 2.5% over the study period.
234
235 Among seroincident cases between round 15 and 18, sequence data for NNRTI resistance
236 substitutions was available for 223 PLHIV (69.69%). We estimated that the probability of
237 harboring an NNRTI-resistant genotype among seroincident PLHIV was 9.95% (95% CI 4.65% -
238 21.28%) in round 18, which was similar to the probability in round 15 through 17 (appendix p
239 37). This analysis was underpowered to detect the difference in the prevalence of NNRTI-
240 resistant genotypes observed in the population-level analysis (appendix p 38).
241
242 Among viremic treatment-experienced round 18 participants, NNRTI and NRTI resistant
243 genotypes were identified in 47.67 % (95 CI 40.94%-55.50%) and 36.55% (95% CI 30.14%-
244 44.31%), respectively, which was similar to the observed prevalence in round 17 (figure 1E,
245 appendix p 39). Only 2.18% (95% CI 0.83%-5.76%) had viruses with PI resistance.
246
247 *Multi-class resistance among PLHIV*
248 Among viremic pre-treatment PLHIV with genotype-data available for all NNRTI, NRTI, and
249 PIs and resistance to at least one of these drug classes (N=236), 83.05% harbored single-class
250 resistance, of which the majority was NNRTI single-class resistance (figure 2A). In contrast,

among treatment-experienced participants (N=197), less than a third had viruses with single-class resistance and 72.59% harbored NNRTI/NRTI (/PI) multi-class resistant genotypes (figure 2B). In round 18 specifically, the prevalence of NNRTI multi-class resistance among viremic PLHIV was estimated to be 1.37-folder higher than the prevalence of NNRTI single-class resistance (figure 2C, appendix pp 40-41).

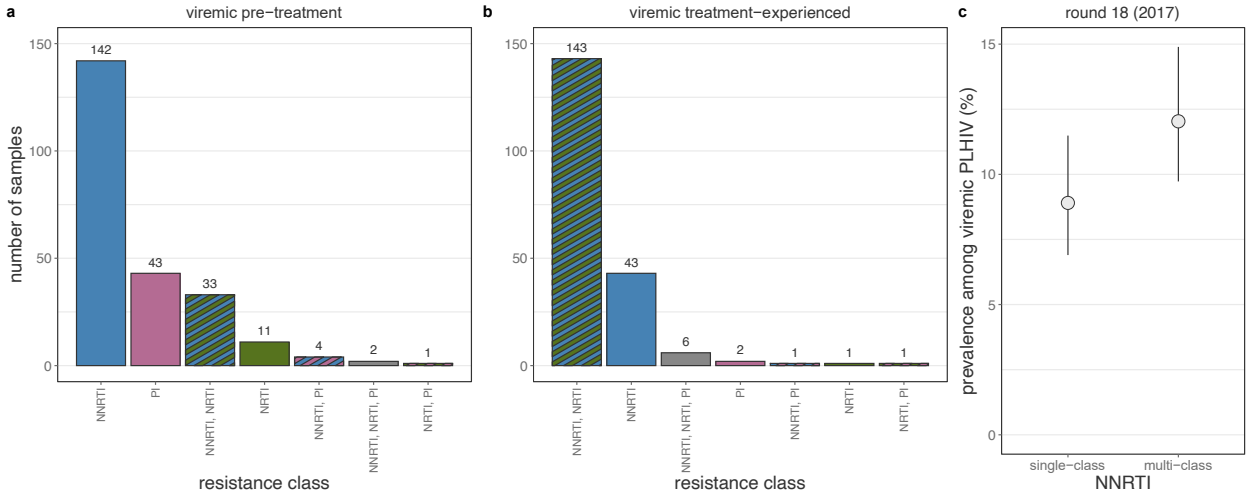


Figure 2: Patterns of multi-class resistance in Rakai Community Cohort Study. (A) Multi-class resistance among 236 viremic pre-treatment PLHIV with genotype data for all NNRTIs, NRTIs, PIs, and resistance to at least one of these drug classes. (B) Multi-class resistance among 197 viremic treatment-experienced PLHIV with genotype data for all NNRTIs, NRTIs, PIs, and resistance to at least one of these drug classes. Counts of each category are shown above the bar. (C) Estimated prevalence of NNRTI single-class and multi-class resistance in round 18 (2017) among viremic PLHIV. Estimates were generated using robust log-Poisson regression with robust standard errors. Error bars indicate the Wald 95% confidence interval for the mean value within each category. Only participant-visits with sequence data for all NNRTIs, NRTIs, and PIs were included. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = integrase inhibitors.

Geographic heterogeneity in resistance

We further stratified our estimates of the population prevalence of viremic resistance in round 18 by community type (agrarian, fishing, and trading) and individual community. The prevalence of NNRTI and NRTI resistance was higher in fishing communities (2.35%, 95% CI 1.86% - 2.97% and 1.46%, 95% CI 1.08% - 1.96%, respectively) compared to trading (0.59%, 95% CI 0.41% - 0.85% and 0.27%, 95% CI 0.16% - 0.46% or agrarian (0.49%, 95% CI 0.35% - 0.69% and 0.27% 95% CI 0.18% - 0.43%) communities (figure 3, appendix p 42).

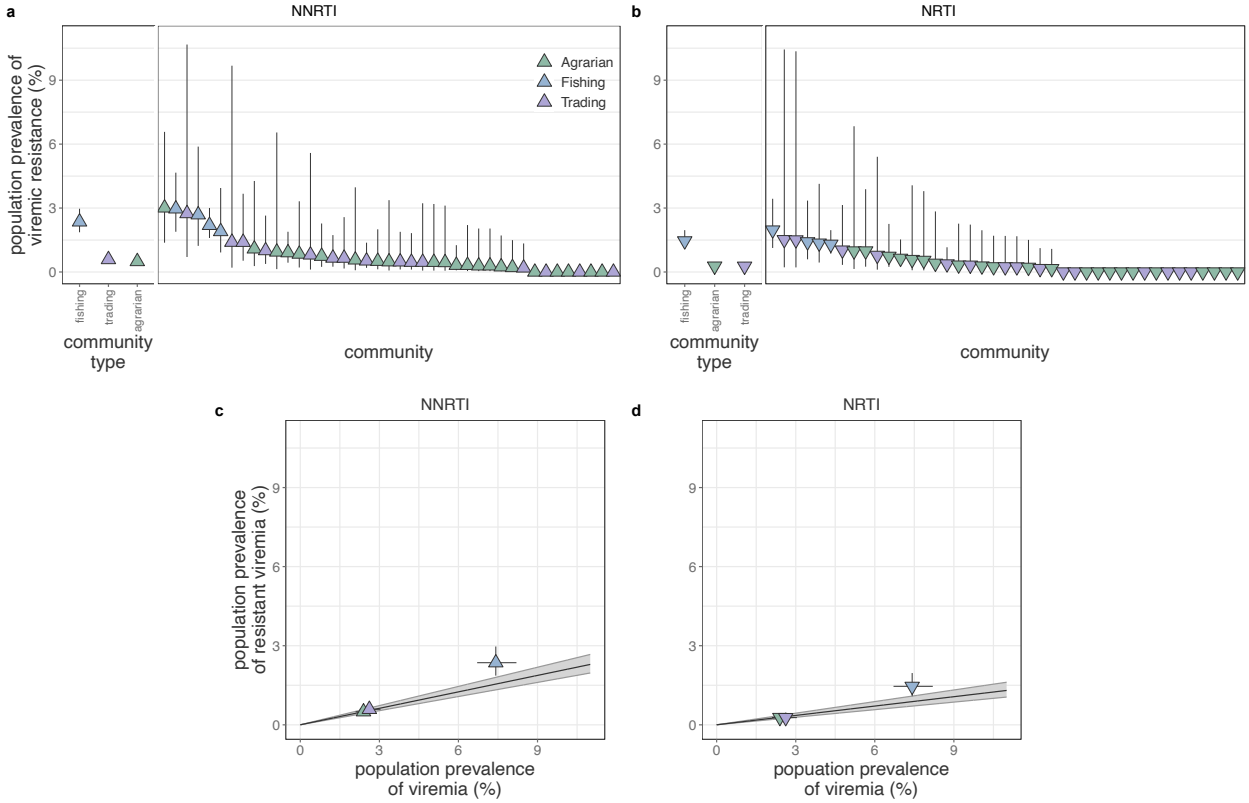


Figure 3: Prevalence of NNRTI and NRTI resistant genotypes in the Rakai Community Cohort Study survey round 18 stratified by community. Estimated prevalence of viremic resistance to NNRTIs (A) and NRTIs (B) stratified by community type (agrarian shown in green, fishing shown in blue, and trading shown in purple), and community (colored by type, unlabeled). Vertical bars represent the 95% confidence intervals for the estimated mean value for each category. (C,D) Estimated prevalence of viremic resistance (as shown in A and B) as a function of the estimated prevalence of viremia by drug class (C, NNRTI; D, NRTI) for each community type. Community type colored as in A, B, and C. Background line and shaded region represent the expected prevalence of resistant viremia and 95% confidence intervals as a function of the prevalence of viremia estimated using the mean prevalence of resistance among all viremic individuals, regardless of community. Estimates were generated using robust log-Poisson regression with 95% Wald confidence intervals around the mean value. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors.

Further, among fishing communities, the population prevalence of NNRTI and NRTI resistance was greater than expected based on the prevalence of viremia in fishing communities and the average prevalence of resistant genotypes among all viremic PLHIV regardless of community (figure 3B,C, appendix p 42).

Prevalence of amino acid substitutions associated with drug resistance

Among participant-visits with genotyping data available for all drugs, the most common drug resistance substitutions in the reverse transcriptase gene (those giving rise to NNRTI and NRTI resistance) were rtK103N (NNRTI, 8.82%), rtM184V (NRTI, 6.93%), and rtE138A (NNRTI and NRTI, 4.32%, figure 4A). Among viremic pre-treatment PLHIV, the prevalence of rtK103N was 7.27% (95% CI 5.17% - 10.20%) in round 18, a 2.36-fold (95% CI 1.13 – 4.92) increase from round 15 (figure 4B,C, appendix pp 45-46). We observed no difference in the pre-treatment prevalence of rtK103N between men and women (appendix p 47). Unsurprisingly, rtK103N was more common among viremic treatment-experienced individuals (26.73%, 95% CI 20.93% -

34.14%, figure 4D, appendix p 48). However, rtM184V was the most common substitution in this population and was present in approximately a third of treatment-experienced individuals.

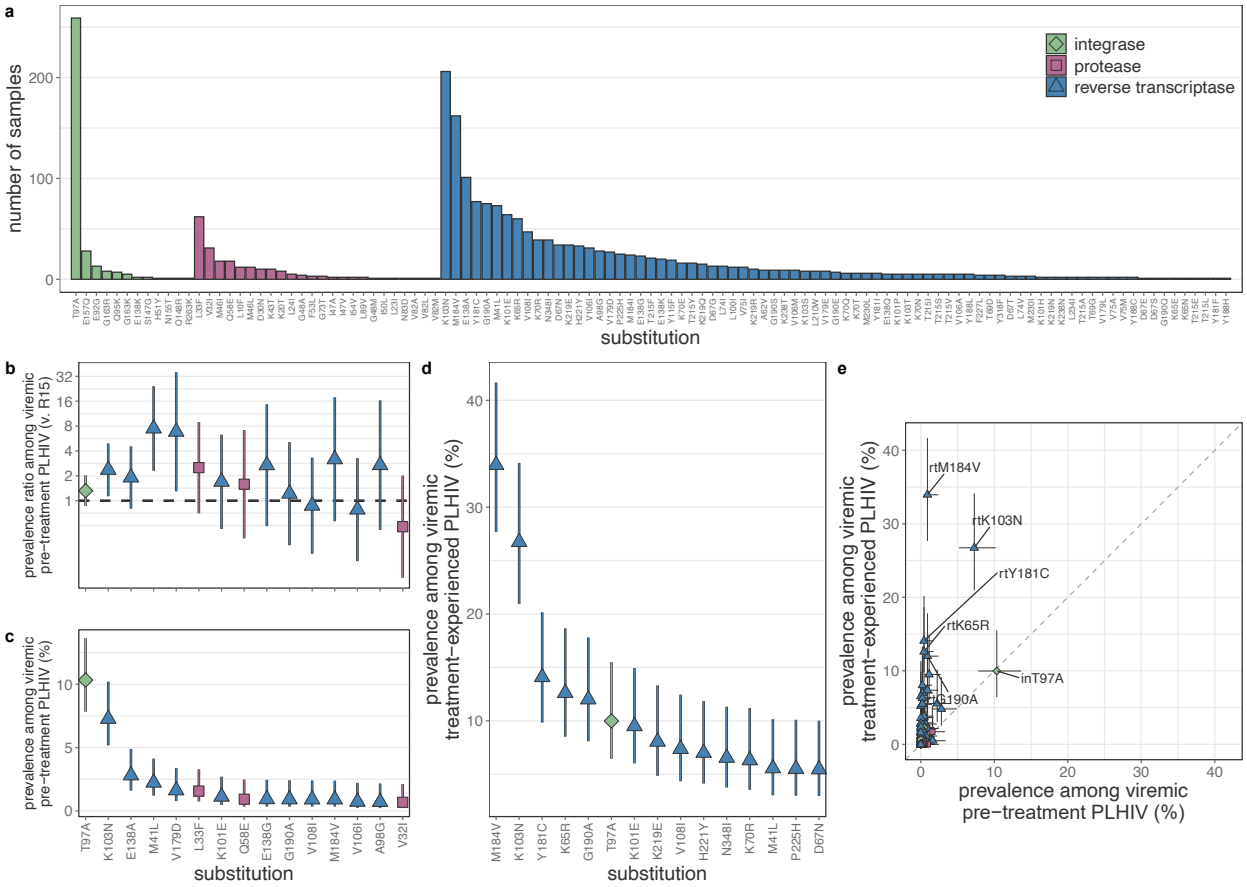


Figure 4: Amino acid substitutions underlying predicted drug resistance. (A) Number of sequenced participant visits (N=4,130) in which each drug resistance substitution is observed. Data points are colored based on the gene in which each substitution occurs (integrase shown in green, protease shown in pink, and reverse transcriptase shown in blue). Within gene substitutions are ordered based on the number of observations. (B) Prevalence ratio comparing survey round 18 to survey round 15 of the 15 most prevalent drug resistance substitutions in pre-treatment PLHIV, sorted by round 18 prevalence. (C) Prevalence in survey round 18 of the 15 most prevalent substitutions in viremic pre-treatment PLHIV sorted by prevalence. (D) Prevalence in survey round 18 of the 15 most prevalent substitutions in treatment-experienced PLHIV sorted by prevalence. (E) Prevalence of all observed resistant substitutions among viremic pre-treatment PLHIV (X-axis) and viremic treatment-experienced PLHIV (Y-axis). Select substitutions are labelled. Estimates generated using robust log-Poisson regression. The Wald method was used to generate 95% confidence intervals around the mean value. Data points are colored by gene as in (A) and genes are further indicated by point shapes (integrase shown as a diamond, protease as a square, and reverse transcriptase as a triangle). PLHIV = people living with HIV. RT = reverse transcriptase. IN = integrase. PR = protease.

Overwhelmingly, the most frequent resistance associated substitution was inT97A (11.09% of viremic genotyped participant-visits, figure 4A). This substitution is canonically associated with integrase strand transfer inhibitors (INSTI) resistance, although its impacts in the absence of other INSTI mutations are thought to be minimal.[44] Overall this substitution was present in 10.33% (95% CI 7.83% - 13.63%) of viremic pre-treatment PLHIV in round 18, which was similar to its prevalence in round 15 (PR 1.32, 95% CI 0.86 – 2.02, figure 5B,C). In contrast to other drug resistance substitutions, the prevalence of inT97A was nearly identical among viremic pre-treatment and treatment-experienced individuals. The estimated prevalence of inT97A

among viremic pre-treatment PLHIV was robust to the assumed correlation structure between repeated measures and observed irrespective of sex, age, community type, or HIV-1 group M subtype (appendix pp 49-53). We did not observe an increased viral load among pre-treatment individuals with in97A as opposed to in97T (appendix p 54). Finally, within-host frequency and number of deduplicated supporting reads were comparable between T97A and all other resistance associated substitutions.

Discussion

Here, we report on trends in the prevalence of HIV drug resistance among viremic individuals in a longitudinal population-based open cohort study in southern Uganda during a period of intensive scale-up of antiretroviral treatment programs. HIV deep-sequence data were available for survey round 15 (2012) through 18 (2017) and partially for survey round 19 (2019). Due to increased treatment uptake and significantly lower HIV incidence, the population prevalence of pre-treatment NNRTI and NRTI resistance decreased by 62% and 80% among all participants regardless of HIV status, respectively, between survey round 15 (2012) and 18 (2017). In survey round 18, 0.25% and 0.04% of all participant visits were contributed by viremic pre-treatment PLHIV with NNRTI and NRTI resistant viremia. However, this decrease was slower than that of susceptible viremia and the prevalence of NNRTI resistance approximately doubled among pre-treatment PLHIV over this time period, reaching nearly 10% by round 18. Resistance was more prevalent among viremic treatment-experienced PLHIV of whom 48% and 37% harbored NNRTI and NRTI resistance (round 18). We did not observe a decrease in the prevalence of resistance among viremic treatment-experienced PLHIV between round 17 and 18. Despite a lack of INSTI use in this cohort over the study period, approximately 10% of viremic pre-treatment and treatment-experienced PLHIV harbored the inT97A substitution.

The estimated increase in NNRTI resistance among viremic pre-treatment PLHIV from 5% in survey round 15 to 10% in round 18 is consistent with a cross-sectional meta-analysis of predominantly clinic-based studies conducted in eastern Africa between 2017 and 2022[11] and with the longitudinal prevalence estimates reported in a meta-analysis of cross-sectional studies from eastern Africa.[10] Our estimated prevalence of pre-treatment NNRTI resistance is slightly less than what was observed in a recent analysis of 291 care-seeking PLHIV in Uganda,[12] which likely reflects differences in the burden of resistance between those who are, and aren't engaged in care and highlights the value in estimates from population-based cohort studies.

Among pre-treatment PLHIV, NNRTI resistance was largely driven by the rtK103N substitution (efavirenz and nevirapine resistance[45]) which increased in prevalence six-fold over the study period and was present in 7% and 26% of viremic pre-treatment and treatment-experienced PLHIV during survey round 18. The high prevalence of rtK103N among pre-treatment PLHIV indicates that this substitution can persist without reversion in the absence of treatment, consistent with a limited fitness cost.[46] NNRTI resistance in this cohort may be partially attributable to single-dose nevirapine treatment of pregnant women used to prevent mother-to-child-transmission in Uganda between 2000 and 2006[47] which has been shown to increase resistance in women and children.[48]

Approximately half of participants who remained viremic after treatment initiation in this study were not found to harbor resistance. This suggests that first-line regimens may remain effective

376 in a subset of this population and additional adherence counseling may be needed to increase
377 levels of viral suppression.[49]
378
379 Among treatment-experienced individuals with resistant genotypes, most harbored NNRTI/NRTI
380 multi-class resistance, followed by NNRTI single-class resistance. This suggests that within
381 individuals NNRTI resistance arises first followed by resistance to NRTI, consistent with
382 previous work.[3,50,51] In contrast, NNRTI mono-resistance dominated among pre-treatment
383 PLHIV. As the prevalence of multi-class resistance among viremic PLHIV was approximately
384 1.35-times that of single-class resistance, the risk of exposure to multi-class resistance is
385 expected to be greater than that of mono-resistance. Subsequently, a subset of pre-treatment
386 mono-resistant PLHIV may have been infected by a multi-class resistant genotype which
387 reverted to wild-type or alternatively NNRTI/NRTI resistant virus may be less transmissible than
388 NNRTI mono-resistant virus. The former hypothesis would imply that archived NRTI resistance
389 is more prevalent than apparent from sequencing viral RNA. This is potentially concerning as
390 archived NRTI resistance could result in de factor mono-therapy under the current first-line
391 NRTI/DTG treatment regimen. Encouragingly, DTG monotherapy has been shown to be robust
392 to resistance evolution for at least 192 weeks.[52] However, treatment adherence was extremely
393 high (>99%) in this cohort and these results may not hold in other settings.
394
395 Approximately a tenth of sequenced participant-visits from recent survey rounds harbored the
396 inT97A substitution. Alone, inT97A has minimal impact on susceptibility to most INSTIs,[44]
397 however, in combination with other substitutions (e.g. inG140S and inQ148H) it can
398 significantly increase resistance.[53–55] The addition of inT97A to a inG140S+inQ148H backbone
399 results in an increase from a three-fold to a 28-fold change in DTG EC_{50} compared to wild-
400 type.[54] Further, there is preliminary evidence that inT97A may be under selection as an HLA-
401 escape mutation.[56]
402
403 Lastly, previous work from our group has highlighted the outsized burden of HIV in fishing
404 communities[14], we here report that these communities also have an unusually high prevalence of
405 NNRTI and NRTI resistance. Fishing communities are known have high rates of immigration of
406 PLHIV[57,58] and recent in-migrants into RCCS communities are less likely to be engaged in care
407 and suppressed as compared to long-term residents.[59] Mobile populations of PLHIV may
408 therefore suffer from difficulty accessing treatment leading to treatment interruptions, which can
409 foster acquired evolution.
410
411 There are some important limitations of this work. Due to unknown HIV sero-status among non-
412 participating residents of RCCS communities, we did not attempt to generalize our results
413 beyond study participants. Younger individuals, men, and residents of trading communities have
414 previously been reported to be less likely to participate in RCCS surveys.[15] Importantly, NNRTI
415 resistance is not significantly associated with these epidemiological sub-groups and we therefore
416 expect participation biases to be minimal. Further, the RCCS only routinely collects self-reported
417 treatment status, which can lead to misclassification of participants. Prior work has demonstrated
418 that 11% of self-reported ART-naïve participants had antivirals in their blood.[60] Given the
419 significant differences observed in the mutational profiles based on self-reported treatment status
420 and the consistency of these results with estimates of the fitness impact of mutations in the
421 absence of treatment, [46] we expect that these biases in the current study are minimal. As this

work is based on deep-sequence data generated from viral RNA in plasma samples, we were only able to identify resistance substitutions among viremic PLHIV. As such, our population prevalence estimates are necessarily an underestimate of the true values as a subset of suppressed PLHIV may harbor resistance if they have successfully transitioned to second-line therapy or are transiently suppressed shortly after initiating treatment. The latter scenario may be more pronounced in recent survey rounds as treatment scale-up has likely increased the number of treatment-experienced individuals who have very recently initiated treatment.

This study adds critical context to our understanding of the HIV epidemic in southern Uganda and to the impact of treatment expansion on the population burden of HIV resistance. While HIV drug resistance increased among pre-treated individuals, increasing VL suppression in the population as a result of treatment was associated with declines in the overall prevalence of viremic resistant HIV in the population.

Our results further suggest that given the relatively high prevalence of viremic treatment-experienced PLHIV with multi-class NNRTI/NRTI resistance, archived NRTI resistance among pre-treatment PLHIV may be more common than is apparent based on sequencing of viral RNA. Future studies that use proviral sequencing to identify archived resistance directly are needed to evaluate this hypothesis. The high-prevalence of inT97A, which appears to be an important compensatory mutation for DTG resistance and has little to no negative fitness impacts in the absence of drug, is concerning. Overall, these findings stress the importance of continued viral sequence-based monitoring of resistance mutations among PLHIV during the ongoing scale-up of DTG and NRTI-based treatment regimens.

Acknowledgements
We thank the Rakai Health Sciences Program and the participants of the Rakai Community Cohort Study for making this research possible. We thank members of the PANGEA-HIV consortium, the Johns Hopkins University Infectious Disease Dynamics group, and the NIAID Laboratory of Immunoregulation International HIV/STD Section at the Johns Hopkins University School of Medicine for helpful feedback on this work. This work was supported by the National Institutes of Health (NIH) National Institute of Allergy and Infectious Diseases (U01AI051171, U01AI075115, UM1AI069530-16, R01AI087409, U01AI100031, R01AI110324, R01AI114438, K25AI114461, R01AI123002, K01AI125086, R01AI128779, R01AI143333, R21AI145682, R01AI155080, ZIAAI001040), NIH National Institute of Child Health and Development (R01HD038883, R01HD050180, R01HD070769, R01HD091003), NIH National Heart, Lung, and Blood Institute (R01HL152813), NIAID Division of Intramural Research, the Fogarty International Center (D43TW009578, D43TW010557), the Johns Hopkins University Center for AIDS Research (P30AI094189), the Bill & Melinda Gates Foundation (OPP1175094, OPP1084362), and the U.S. President's Emergency Plan for AIDS Relief through the Centers for Disease Control and Prevention (NU2GGH000817).

Data sharing
Code to reproduce all analyses and visualizations as well as de-identified resistance and limited patient metadata are available at https://github.com/m-a-martin/rccs_hiv_resistance_r15_r19. Due to privacy concerns we are unable to share individual-level data on community of residence.

468 **Declaration of interests**
469 We declare no competing interests.
470
471 **References**

472 1 Hoenigl M, Chaillon A, Moore DJ, *et al.* Rapid HIV Viral Load Suppression in those
473 Initiating Antiretroviral Therapy at First Visit after HIV Diagnosis. *Sci Rep* 2016; **6**: 32947.

474 2 Kelley CF, Kitchen CMR, Hunt PW, *et al.* Incomplete Peripheral CD4 $^{+}$ Cell Count
475 Restoration in HIV-Infected Patients Receiving Long-Term Antiretroviral Treatment. *Clin*
476 *Infect Dis* 2009; **48**: 787–94.

477 3 Cohen MS, Chen YQ, McCauley M, *et al.* Antiretroviral Therapy for the Prevention of HIV-1
478 Transmission. *N Engl J Med* 2016; **375**: 830–9.

479 4 UNAIDS. AIDSinfo. https://aidsinfo.unaids.org.

480 5 Sethi AK, Celentano DD, Gange SJ, Moore RD, Gallant JE. Association between Adherence
481 to Antiretroviral Therapy and Human Immunodeficiency Virus Drug Resistance. *Clin Infect*
482 *Dis* 2003; **37**: 1112–8.

483 6 Bangsberg DR, Charlebois ED, Grant RM, *et al.* High levels of adherence do not prevent
484 accumulation of HIV drug resistance mutations: *AIDS* 2003; **17**: 1925–32.

485 7 Boender TS, Sigaloff KCE, McMahon JH, *et al.* Long-term Virological Outcomes of First-
486 Line Antiretroviral Therapy for HIV-1 in Low- and Middle-Income Countries: A Systematic
487 Review and Meta-analysis. *Clin Infect Dis* 2015; **61**: 1453–61.

488 8 Gregson J, Tang M, Ndembi N, *et al.* Global epidemiology of drug resistance after failure of
489 WHO recommended first-line regimens for adult HIV-1 infection: a multicentre retrospective
490 cohort study. *Lancet Infect Dis* 2016; **16**: 565–75.

491 9 Beck IA, Levine M, McGrath CJ, *et al.* Pre-treatment HIV-drug resistance associated with
492 virologic outcome of first-line NNRTI-antiretroviral therapy: A cohort study in Kenya.
493 *eClinicalMedicine* 2020; **18**: 100239.

494 10 Gupta RK, Gregson J, Parkin N, *et al.* HIV-1 drug resistance before initiation or re-initiation
495 of first-line antiretroviral therapy in low-income and middle-income countries: a systematic
496 review and meta-regression analysis. *Lancet Infect Dis* 2018; **18**: 346–55.

497 11 Ntamatungiro AJ, Kagura J, Weisser M, Francis JM. Pre-treatment HIV-1 drug resistance in
498 antiretroviral therapy-naive adults in Eastern Africa: a systematic review and meta-analysis. *J*
499 *Antimicrob Chemother* 2022; **77**: 3231–41.

500 12 Crowell TA, Danboise B, Parikh A, *et al.* Pretreatment and Acquired Antiretroviral Drug
501 Resistance Among Persons Living With HIV in Four African Countries. *Clin Infect Dis* 2021;
502 **73**: e2311–22.

503 13 Consolidated guidelines for the prevention and treatment of HIV in Uganda. 2018; published
504 online Sept.

505 14 Chang LW, Grabowski MK, Ssekubugu R, *et al.* Heterogeneity of the HIV epidemic in
506 agrarian, trading, and fishing communities in Rakai, Uganda: an observational
507 epidemiological study. *Lancet HIV* 2016; **3**: e388–96.

508 15 Grabowski MK, Serwadda DM, Gray RH, *et al.* HIV Prevention Efforts and Incidence of HIV
509 in Uganda. *N Engl J Med* 2017; **377**: 2154–66.

510 16 Kagulire SC, Opendi P, Stamper PD, *et al.* Field evaluation of five rapid diagnostic tests for
511 screening of HIV-1 infections in rural Rakai, Uganda. *Int J STD AIDS* 2011; **22**: 308–9.

512 17 Elm EV, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening
513 the reporting of observational studies in epidemiology (STROBE) statement: guidelines for
514 reporting observational studies. *BMJ* 2007; **335**: 806–8.

515 18 Pillay D, Herbeck J, Cohen MS, *et al.* PANGEA-HIV: phylogenetics for generalised
516 epidemics in Africa. *Lancet Infect Dis* 2015; **15**: 259–61.

517 19 Abeler-Dörner L, Grabowski MK, Rambaut A, Pillay D, Fraser C. PANGEA-HIV 2:
518 Phylogenetics And Networks for Generalised Epidemics in Africa. *Curr Opin HIV AIDS*
519 2019; **14**: 173–80.

520 20 Monod M, Brizzi A, Ssekubugu R, *et al.* Growing gender disparity in HIV infection in Africa:
521 sources and policy implications. .

522 21 Bonsall D, Golubchik T, de Cesare M, *et al.* A Comprehensive Genomics Solution for HIV
523 Surveillance and Clinical Monitoring in Low-Income Settings. *J Clin Microbiol* 2020; **58**:
524 e00382-20.

525 22 Wymant C, Blanquart F, Golubchik T, *et al.* Easy and accurate reconstruction of whole HIV
526 genomes from short-read sequence data with shiver. *Virus Evol* 2018; **4**.
527 DOI:10.1093/ve/vey007.

528 23 Siepel AC, Halpern AL, Macken C, Korber BTM. A Computer Program Designed to Screen
529 Rapidly for HIV Type 1 Intersubtype Recombinant Sequences. *AIDS Res Hum Retroviruses*
530 1995; **11**: 1413–6.

531 24 Fogel JM, Bonsall D, Cummings V, *et al.* Performance of a high-throughput next-generation
532 sequencing method for analysis of HIV drug resistance and viral load. *J Antimicrob*
533 *Chemother* 2020; **75**: 3510–6.

534 25 Broad Institute. Picard tools. Https://Broadinstitute.Github.Io/Picard/. 2016.
535 https://broadinstitute.github.io/picard/%5Cnhttp://broadinstitute.github.io/picard/.

536 26 McGinnis S, Madden TL. BLAST: at the core of a powerful and diverse set of sequence
537 analysis tools. *Nucleic Acids Res* 2004; **32**: W20–5.

538 27 Ji H, Enns E, Brumme CJ, *et al.* Bioinformatic data processing pipelines in support of next-
539 generation sequencing-based HIV drug resistance testing: the Winnipeg Consensus. *J Int*
540 *AIDS Soc* 2018; **21**: e25193.

541 28 Rhee S-Y. Human immunodeficiency virus reverse transcriptase and protease sequence
542 database. *Nucleic Acids Res* 2003; **31**: 298–303.

543 29 Shafer RW. Rationale and Uses of a Public HIV Drug-Resistance Database. *J Infect Dis* 2006;
544 **194**: S51–8.

545 30 Liu TF, Shafer RW. Web Resources for HIV Type 1 Genotypic-Resistance Test
546 Interpretation. *Clin Infect Dis* 2006; **42**: 1608–18.

547 31 R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Australia,
548 2020 https://www.R-project.org/ (accessed Dec 11, 2020).

549 32 Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data.
550 *Am J Epidemiol* 2004; **159**: 702–6.

551 33 Zeileis A. Object-Oriented Computation of Sandwich Estimators. *J Stat Softw* 2006; **16**.
552 DOI:10.18637/jss.v016.i09.

553 34 Zeileis A, Hothorn T. Diagnostic Checking in Regression Relationships. *R News* 2002; **2**: 7–
554 10.

555 35 Halekoh U, Højsgaard S, Yan J. The *R* Package **geepack** for Generalized Estimating
556 Equations. *J Stat Softw* 2006; **15**. DOI:10.18637/jss.v015.i02.

557 36 Wickham H, Averick M, Bryan J, *et al.* Welcome to the Tidyverse. *J Open Source Softw*
558 2019; **4**: 1686.

559 37 Wickham H, Chang W, Henry L, *et al.* ggplot2: Elegant Graphics for Data Analysis. Springer-
560 Verlag New York, 2021 http://ggplot2.org (accessed Feb 8, 2022).

561 38 Wilke CO. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. 2020.
562 https://CRAN.R-project.org/package=cowplot.

563 39 Lin Pedersen T. patchwork: The Composer of Plots. 2023. https://cran.r-
564 project.org/web/packages/patchwork/index.html.

565 40 FC M, Davis TL, ggplot2 authors. ggpattern: 'ggplot2' Pattern Geoms. 2022.
566 https://github.com/coolbutuseless/ggpattern.

567 41 Slowikowski K. ggrepel: Automaticalyl Pposition Non-Overlapping Text Labels with
568 'ggplot2'. 2023. https://github.com/slowkow/ggrepel.

569 42 Wickham H, Bryan J. readxl: Read Excel Files. 2023. https://readxl.tidyverse.org.

570 43 Wickham H, Miller E, Smith D. haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files.
571 2023. https://github.com/tidyverse/haven.

572 44 Margot NA, Ram RR, White KL, Abram ME, Callebaut C. Antiviral activity of HIV-1
573 integrase strand-transfer inhibitors against mutants with integrase resistance-associated
574 mutations and their frequency in treatment-naïve individuals. *J Med Virol* 2019; **91**: 2188–94.

575 45 Melikian GL, Rhee S-Y, Varghese V, *et al.* Non-nucleoside reverse transcriptase inhibitor
576 (NNRTI) cross-resistance: implications for preclinical evaluation of novel NNRTIs and
577 clinical genotypic resistance testing. *J Antimicrob Chemother* 2014; **69**: 12–20.

578 46 Kühnert D, Kouyos R, Shirreff G, *et al.* Quantifying the fitness cost of HIV-1 drug resistance
579 mutations through phylodynamics. *PLOS Pathog* 2018; **14**: e1006895.

580 47 Namara-Lugolobi E, Namukwaya Z, Owor M, *et al.* Twenty years of Prevention of Mother to
581 Child HIV Transmission: research to implementation at a national referral hospital in Uganda.
582 *Afr Health Sci* 2022; **22**. DOI:10.4314/ahs.v22i2.5S.

583 48 Arrivé E, Newell M-L, Ekouevi DK, *et al.* Prevalence of resistance to nevirapine in mothers
584 and children after single-dose exposure to prevent vertical transmission of HIV-1: a meta-
585 analysis†. *Int J Epidemiol* 2007; **36**: 1009–21.

586 49 Chang LW, Serwadda D, Quinn TC, Wawer MJ, Gray RH, Reynolds SJ. Combination
587 implementation for HIV prevention: moving from clinical trial evidence to population-level
588 effects. *Lancet Infect Dis* 2013; **13**: 65–76.

589 50 Feder AF, Harper KN, Brumme CJ, Pennings PS. Understanding patterns of HIV multi-drug
590 resistance through models of temporal and spatial drug heterogeneity. *eLife* 2021; **10**: e69032.

591 51 Reynolds SJ, Kityo C, Mbamanya F, *et al.* Evolution of drug resistance after virological
592 failure of a first-line highly active antiretroviral therapy regimen in Uganda. *Antivir Ther*
593 2009; **14**: 293–7.

594 52 West E, Zeeb M, Grube C, *et al.* Sustained viral suppression with dolutegravir monotherapy
595 over 192 weeks in patients starting combination antiretroviral therapy during primary HIV
596 infection (EARLY-SIMPLIFIED): a randomized, controlled, multi-site, non-inferiority trial.
597 *Clin Infect Dis* 2023; : ciad366.

598 53 Huik K, Hill S, George J, *et al.* High-level dolutegravir resistance can emerge rapidly from
599 few variants and spread by recombination: implications for integrase strand transfer inhibitor
600 salvage therapy. *AIDS* 2022; **36**: 1835–40.

601 54 Zhang WW, Cheung PK, Oliveira N, Robbins MA, Harrigan PR, Shahid A. Accumulation of
602 Multiple Mutations In Vivo Confers Cross-Resistance to New and Existing Integrase
603 Inhibitors. *J Infect Dis* 2018; **218**: 1773–6.

604 55 George JM, Kuriakose SS, Dee N, *et al.* Rapid Development of High-Level Resistance to
605 Dolutegravir With Emergence of T97A Mutation in 2 Treatment-Experienced Individuals
606 With Baseline Partial Sensitivity to Dolutegravir. *Open Forum Infect Dis* 2018; **5**: ofy221.

607 56 Brumme ZL, John M, Carlson JM, *et al.* HLA-Associated Immune Escape Pathways in HIV-1
608 Subtype B Gag, Pol and Nef Proteins. *PLoS ONE* 2009; **4**: e6687.

609 57 Ratmann O, Kagaayi J, Hall M, *et al.* Quantifying HIV transmission flow between high-
610 prevalence hotspots and surrounding communities: a population-based study in Rakai,
611 Uganda. *Lancet HIV* 2020; **7**: e173–83.

612 58 Kate Grabowski M, Lessler J, Bazaale J, *et al.* Migration, hotspots, and dispersal of HIV
613 infection in Rakai, Uganda. *Nat Commun* 2020; **11**: 976.

614 59 Billioux VG, Chang LW, Reynolds SJ, *et al.* Human immunodeficiency virus care cascade
615 among sub-populations in Rakai, Uganda: an observational study: Billioux VG et al. *J Int*
616 *AIDS Soc* 2017; **20**: 21590.

617 60 Grabowski MK, Reynolds SJ, Kagaayi J, *et al.* The validity of self-reported antiretroviral use
618 in persons living with HIV: a population-based study. *AIDS* 2018; **32**: 363–9.

619
620

Table 1: Total number of participant visits in each survey round stratified by HIV status

	Participant-visits			
		PLHIV		
		All	Viremic	
Round	**All**		**All**	**Pre-treatment**
Overall	93659	17471 (18.65%)	5953 (34.07%)	5177 (86.96%)
Survey round				
15 (August 2011 - May 2013)	17170	3498 (20.37%)	..	2246 (98.47%)
16 (July 2013 - January 2015)	18003	3391 (18.84%)	1465 (43.2%)	1319 (90.03%)
17 (February 2015 - September 2016)	19342	3618 (18.71%)	1026 (28.36%)	822 (80.12%)
18 (October 2016 - May 2018)	19811	3638 (18.36%)	728 (20.01%)	502 (68.96%)
19 (June 2018 - November 2020)	19333	3326 (17.2%)	453 (13.62%)	288 (63.58%)

Dates listed for each survey round represent the range over which survey visits were conducted. Percentages for each category represent the percentage of all participant visits for that round. Overall viremia was not calculated in round 15 due to missing viral load data. PLWHIV = persons living with HIV.

Table 2: Number of viremic samples at participant visits that had successful drug genotyping for at least one drug.

	Viremic PLWHIV				Viremic pre-treatment PLWHIV				Viremic treatment-experienced PLWHIV		
	Participant-visits	**Genotype data for at least one drug class (%)**	***p*-value**		**Participant-visits**	**Genotype data for at least one drug class (%)**	***p*-value**		**Participant-visits**	**Genotype data for at least one drug class (%)**	***p*-value**
Overall	5953	4130 (69.38%)			5177	3713 (71.72%)			776	417 (53.74%)	
Age Category			0.15				0.31				0.82
(14,24]	1146	845 (73.73%)			1032	780 (75.58%)			114	65 (57.02%)	
(24,34]	2931	2048 (69.87%)			2573	1852 (71.98%)			358	196 (54.75%)	
(34,50]	1876	1237 (65.94%)			1572	1081 (68.77%)			304	156 (51.32%)	
Community type			0.61				0.81				0.17
Agrarian	1794	1212 (67.56%)			1590	1117 (70.25%)			204	95 (46.57%)	
Fishing	2861	2024 (70.74%)			2455	1782 (72.59%)			406	242 (59.61%)	
Trading	1298	894 (68.88%)			1132	814 (71.91%)			166	80 (48.19%)	
Sex			0.76				0.88				0.95
F	3188	2198 (68.95%)			2744	1961 (71.47%)			444	237 (53.38%)	
M	2765	1932 (69.87%)			2433	1752 (72.01%)			332	180 (54.22%)	
Survey round			<1e-4				<1e-4				<1e-4
15 (2012)	2281	1750 (76.72%)			2246	1749 (77.87%)			35	1 (2.86%)	
16 (2014)	1465	658 (44.91%)			1319	654 (49.58%)			146	4 (2.74%)	
17 (2015)	1026	921 (89.77%)			822	744 (90.51%)			204	177 (86.76%)	
18 (2017)	728	631 (86.68%)			502	444 (88.45%)			226	187 (82.74%)	
19 (2019)	453	170 (37.53%)			288	122 (42.36%)			165	48 (29.09%)	
Viral load (log10 copies/mL)			0.084				0.063				0.95
(3,4]	2101	1360 (64.73%)			1748	1168 (66.82%)			353	192 (54.39%)	
(4,5]	2231	1631 (73.11%)			1898	1452 (76.5%)			333	179 (53.75%)	
(5,Inf]	598	421 (70.4%)			508	375 (73.82%)			90	46 (51.11%)	
missing	1023	718 (70.19%)			1023	718 (70.19%)					

Percentages show the percentage of all samples within a category that have successful drug resistance genotyping to at least one drug class. *p*-values generated using the χ^2 test. PLWHIV=people living with HIV.

Table 3: Estimated prevalence of resistance among viremic pre-treatment PLWHIV

	NNRTI				NRTI				PI		
Survey round	**Prevalence (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**		**Prevalence (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**		**Prevalence (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**
15 (2012)	4.97 (3.72, 6.65)	ref	ref		1.68 (1.07, 2.62)	ref	ref		2.1 (1.45, 3.04)	ref	ref
16 (2014)	7.35 (5.04, 10.71)	1.48 (0.92, 2.38)	0.11		1.48 (0.62, 3.57)	0.88 (0.33, 2.37)	0.81		1.9 (0.9, 4.01)	0.9 (0.39, 2.08)	0.81
17 (2015)	7.15 (5.47, 9.35)	1.44 (0.97, 2.14)	0.072		1.85 (1.08, 3.19)	1.11 (0.55, 2.23)	0.78		1.32 (0.69, 2.53)	0.63 (0.3, 1.33)	0.22
18 (2017)	9.77 (7.35, 12.97)	1.96 (1.31, 2.95)	0.0011		1.76 (0.89, 3.51)	1.05 (0.46, 2.39)	0.9		2.02 (1.06, 3.86)	0.96 (0.46, 2.03)	0.92
19 (2019)	8.77 (4.81, 15.99)	1.76 (0.91, 3.44)	0.095		1.65 (0.41, 6.62)	0.98 (0.23, 4.23)	0.98		0.83 (0.12, 5.81)	0.39 (0.05, 2.86)	0.36

Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. *p*-values that coefficients are different from 0 at the $\alpha = 0.05$ level were calculated using the Wald method. PLWHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = protease inhibitors.

Appendix

Table of Contents

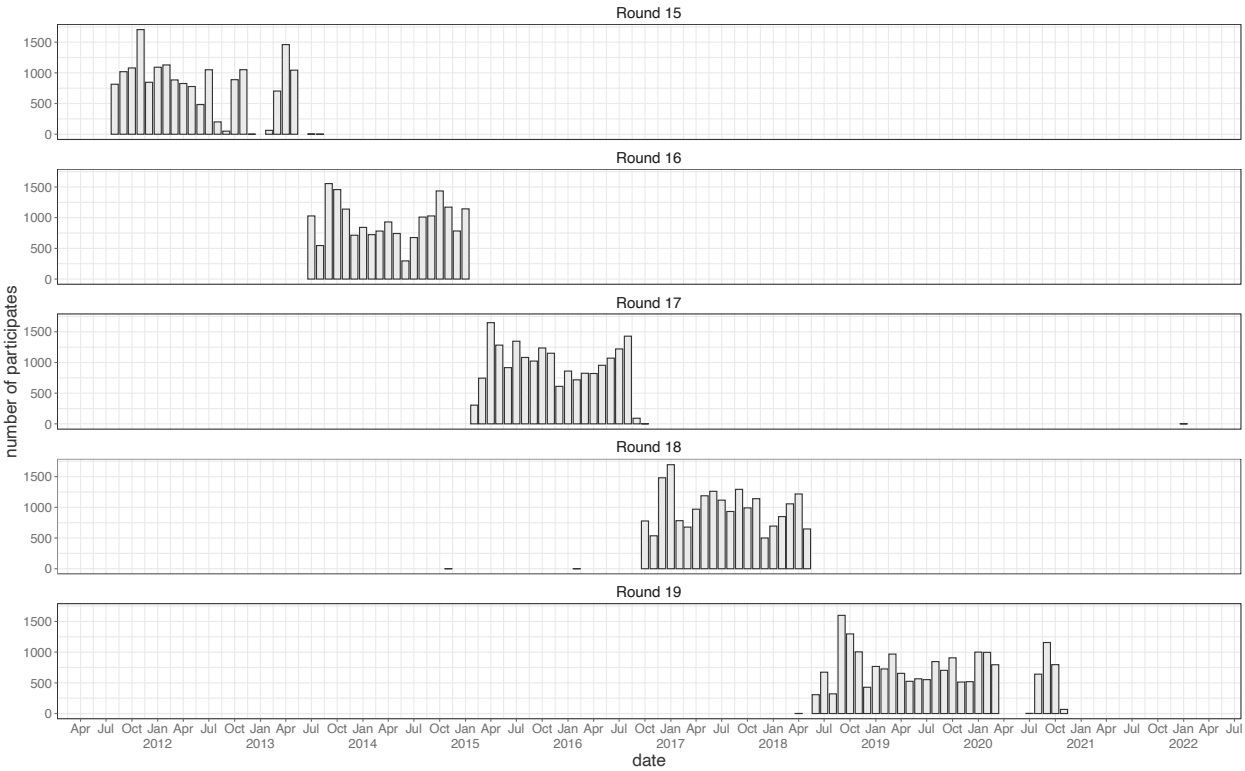
P1…Table of contents
P2…Supplementary Table 1: Rakai Community Cohort Study survey dates
P3…Supplementary Figure 1: Rakai Community Cohort Study interview dates
P4-7…Supplementary methods
P7…Supplementary Table 2: Technology used to generate sequence data by survey round
P8… Supplementary Table 3: Drugs to which resistance was predicted
P9-10… Supplementary Table 4: Summary of outcome measures
P11…Supplementary Table 5: Inverse probability weighting model for the probability that genotyping data is available among all viremic PLHIV
P12-13… Supplementary Table 6: Inverse probability weighting model for the probability that genotyping data is available among viremic pre-treatment PLHIV
P14… Supplementary Table 7: Inverse probability weighting model for the probability that genotyping data is available among viremic treatment-experienced PLHIV
P15…Supplementary Table 8: Inverse probability weighting model for the probability that genotyping data is available among viremic pre-treatment seroincident PLHIV
P16… Supplementary Figure 2: Number of participant-visits with resistant predictions for each participant
P17…Supplementary Table 9: Prevalence of resistance among viremic PLHIV correlation structure model comparison
P18…Supplementary Table 10: Prevalence of resistance among viremic pre-treatment PLHIV by correlation structure
P19…Supplementary Table 11: Prevalence of resistance among viremic treatment-experienced PLHIV by correlation structure
P20…Supplementary Table 12: Prevalence of resistance among viremic PLHIV by correlation structure
P21…Supplementary Table 13: Prevalence of PLHIV, viremic PLHIV, viremic pre-treatment PLHIV, and viremic treatment-experienced PLHIV
P22…Supplementary Table 14: Number of viremic samples at participant visits that had successful genotyping for all INSTIs
P23…Supplementary Table 15: Number of viremic samples at participant visits that had successful genotyping for all NNRTIs
P24…Supplementary Table 16: Number of viremic samples at participant visits that had successful genotyping for all NRTIs
P25…Supplementary Table 17: Number of viremic samples at participant visits that had successful genotyping for all PIs
P26…Supplementary Table 18: Number of viremic samples that had resistant genotypes for all drug classes
P27…Supplementary Table 19: Number of viremic samples at participant visits that had successful genotyping for all NNRTIs, NRTIs, and PIs
P28…Supplementary Table 20: Number of viremic samples at participant visits that had successful genotyping for all drug classes
P29…Supplementary Table 21: Population prevalence of viremic pre-treatment NNRTI, NRTI, and PI resistance
P30…Supplementary Figure 3: Prevalence of NNRTI, NRTI, and PI resistance among viremic PLHIV in survey round 18, stratified by treatment exposure
P31…Supplementary Table 22: Population prevalence of viremic treatment-experienced NNRTI, NRTI, and PI resistance
P32…Supplementary Table 23: Population prevalence of viremic NNRTI, NRTI, and PI resistance
P33…Supplementary Figure 4: Prevalence of NNRTI resistance among viremic pre-treatment PLHIV stratified by sex, age category, and community type
P34-35…Supplementary Table 24: Prevalence of NNRTI resistance among viremic pre-treatment PLHIV stratified by sex, age category, and community type
P36…Supplementary Table 25: Prevalence of NNRTI resistance among seroincident, pre-treatment PLHIV
P37…Seroincident sub-analysis power calculation
P38…Supplementary Table 26: Prevalence of NNRTI, NRTI, and PI resistance among viremic treatment-experienced PLHIV
P39…Supplementary Figure 5: Patterns of multi-class resistance in Rakai Community Cohort Study, 2017
P40…Supplementary Table 27: Prevalence of NNRTI single- and multi-class resistance among PLHIV
P41…Supplementary Table 28: Population prevalence of NNRTI, NRTI, and PI resistance in survey round 18 stratified by community
P42-43…Supplementary Table 29: Population prevalence of NNRTI, NRTI, and PI resistance in survey round 18 stratified by community
P44-45…Supplementary Table 30: Prevalence of amino acid substitutions among viremic pre-treatment PLHIV
P46…Supplementary Table 31: Prevalence of rtK103N among viremic pre-treatment PLHIV stratified by sex
P47…Supplementary Table 32: Prevalence of amino acid substitutions among viremic treatment-experienced PLHIV
P48… Supplementary Table 33: Prevalence of inT97A among viremic pre-treatment PLHIV by correlation structure
P49…Supplementary Figure 6: Prevalence of inT97A among viremic pre-treatment PLHIV stratified by sex, age category, community type, and subtype
P50-51…Supplementary Table 34: Prevalence of inT97A among viremic pre-treatment PLHIV stratified by sex, age category, community type, and subtype
P52…Supplementary Figure 7: Subtype of viremic pre-treatment PLHIV stratified by presence/absence of inT97A
P53…Supplementary Figure 8: Virological characteristics of viremic pre-treatment PLHIV stratified by presence/absence of inT97A
P54…Supplementary references

Supplementary Table 1: Rakai Community Cohort Study survey dates

Round	Survey start date	Survey mid date	Survey end date
1	1994-11-05	1995-03-20	1995-07-18
2	1995-01-12	1996-01-10	1996-05-01
3	1996-01-12	1996-11-20	1997-04-23
4	1997-06-06	1997-11-21	1998-05-12
5	1997-08-05	1998-08-06	1999-03-15
6	1999-04-06	1999-09-07	2000-02-09
7	2000-03-20	2000-08-08	2001-02-19
8	2001-04-03	2001-10-23	2002-05-31
9	2002-07-15	2003-01-28	2003-08-01
10	2003-09-26	2004-04-23	2004-11-23
11	2005-01-13	2005-10-13	2006-06-30
12	2006-03-31	2007-06-28	2008-04-24
13	2008-06-17	2009-03-17	2009-12-04
14	2010-01-18	2010-09-24	2011-06-21
15	2011-08-10	2012-04-02	2013-05-29
16	2013-07-08	2014-04-07	2015-01-28
17	2015-02-23	2015-11-03	2016-09-02
18	2016-10-03	2017-07-17	2018-05-21
19	2018-06-19	2019-06-20	2020-11-04

Start, mid (median), and end dates for each RCCS survey round. Viral load and sequencing data used in this study were collected during R15-19. R19 sequence data was generated for participant visits through 2019-05-17.

Supplementary Figure 1: Rakai Community Cohort Study interview dates



Number of interviews conducted each month, stratified by Rakai Community Cohort Study round.

Supplementary Methods

HIV deep sequencing

For samples taken at participant visits during survey rounds 15 and 16 RNA extraction was performed at University College London Hospital (London, United Kingdom) using the QIAsymphony SP workstation with the QIAsymphony DSP Virus/Pathogen Kit (Qiagen, Hilden, Germany). Following gel electrophoresis confirmation of amplification samples were sequenced at the Wellcome Trust Sanger Institute (Hinxton, United Kingdom) on Illumina MiSeq and HiSeq platforms.

For the remaining survey rounds, RNA extraction was performed as above at the Oxford Genomics Centre (Oxford, United Kingdom) followed by library preparation with the SMARTer Stranded Total RNA-Seq kit v2 – Pico Input Mammalian (Clontech, TakaRa Bio). Streptavidin-conjugated beads were used to isolate fragments shorter than 400 nucleotides and PCR amplification and purification was performed with the Agencourt AMPure XMP (Beckman Coulter) following the veSEQ-HIV protocol.[1]The Oxford Genomics Centre generated 350 to 600 base pair reads using the Illumina NovaSeq 6000 instrument. A subset of poorly-sequenced treatment-naïve samples from rounds 15 and 16 were also sequenced using this protocol (appendix p 7).

Outcome measures

A summary of primary outcome measures and corresponding figure references is provided in table S4 (appendix pp 9-10).

Population prevalence of all PLHIV, viremic PLHIV, viremic pre-treatment PLHIV, and viremic treatment-experienced PLHIV in RCCS

We define the prevalence of all PLHIV within a given round as the percentage of all participant-visits in that round from individuals with a serologically confirmed HIV infection as described in *Methods*. The prevalence of viremic PLHIV within a given survey round is defined as the percentage of all participants with a positive HIV test result and measured viral load ≥1,000 copies/mL at that survey round. Due to a high proportion of missing viral loads in survey round 15, we did not attempt to quantify the prevalence of viremic HIV or treatment experienced viremic HIV during this round. Within each survey round, pre-treatment PLHIV were identified as those who participants with a serologically confirmed HIV infection who self-report never having taken ART within that round and at all previous rounds in which that individual participated in the study. Participant-visits contributed by treatment-experienced PLHIV were identified as those with a serologically confirmed HIV infection who self-reported ever having taken ART at the current survey round round or at any previous round.

Prevalence of viremic pre-treatment resistance

In this study, individuals with serologically confirmed HIV infection who self-reported never having taken ART (at the current or any other prior survey) were classified as pre-treatment PLHIV. Using robust log-Poisson regression we estimated the prevalence of NNRTI, NRTI, and PI resistance among viremic pre-treatment PLHIV stratified by survey round (R15 – R19) using predicted resistance among all viremic pre-treatment participant visits with available resistance prediction to each class. Drug classes with significant temporal trends among pre-treatment PLHIV were further stratified by age category ((14,24], (24,34], (34,50]), sex (F/M), and community type (agrarian/fishing/trading). Using data from all viremic pre-treatment participant visits with available resistance predictions, non-viremic pre-treatment PLHIV, treatment-experienced PLHIV, and participants not living with HIV we further calculated the population prevalence of viremic pre-treatment NNRTI, NRTI, and PI resistance among all RCCS participants irrespective of HIV serostatus. IPW was used to account for missing resistance data among some viremic pre-treatment PLHIV.

Weights were calculated among pre-treatment PLHIV by first fitting a log-Poisson regression model for each drug class (NNRTI, NRTI, PI) to estimate the probability that resistance prediction was available based on availability of a viral load measurement (True/False), $\log_{10}$ copies/mL among samples with viral load measurements, community type (agrarian/fishing/trading), age category ((14,24]/(24,34]/(34,50]), and sex (M/F) stratified by survey round (appendix pp 12-13). Among participant visits from viremic pre-treatment PLHIV with resistance predictions, weights were calculated as the inverse of the modeled probability of having a resistance prediction normalized such that the sum of the weights within a given drug class and survey round was equal to the number of viremic pre-treatment PLHIV in that survey round. When estimating the prevalence of pre-treatment resistance among all participants, people not living with HIV (PNWHIV), treatment-experienced PLHIV, and non-viremic pre-treatment PLHIV were assigned a weight of 1.

Pre-treatment seroincident cases between rounds 15 and 19 were identified through repeat, longitudinal serologic testing. Specifically, at each visit they were classified as participants testing HIV seropositive for the first time and not reporting ART use with an HIV seronegative test result at the prior visit. We used the subset of these individuals with available NNRTI resistance data in their first seropositive round to estimate the prevalence of NNRTI resistance among seroincident PLHIV in each round. The probability of resistance was modelled as a function of seroconversion survey round using robust log-Poisson regression. IPW was used to account for missing sequence data among seroincident PLHIV as described above.

For all pre-treatment participant-visits from seroincident PLHIV we modeled the probability that NNRTI resistance prediction was available for that visit based on the same predictors as above, excluding community type due to a relatively small number of seroincident PLHIV, also using log-Poisson regression. Regression coefficients were used to calculate the probability that resistance prediction was unavailable for each participant-visit based on these predictors (appendix pp 15).

Next, the probability that at least one NNRTI resistance prediction was available for each seroincident PLHIV was calculated as the complement of the product of the probability that no sequence data was available for all of their visits. The inverse of this probability was taken as the weight for each sero-incident individual with available NNRTI resistance prediction. Weights were normalized such that the sum of the weights equaled the total number of seroincident individuals.

Prevalence of all viremic resistance
In survey rounds 17 and 18 we estimated the population prevalence of viremic PLHIV with NNRTI, NRTI, and PI resistant genotypes, regardless of treatment status, using robust log-Poisson regression. Within each drug class models were fit to data from viremic PLHIV with resistance predictions to that drug, non-viremic PLHIV, and participants not living with HIV, again using IPW to account for missing resistance data.

Weights were calculated by modeling the probability of having NNRTI, NRTI, and PI resistance prediction based on the same predictors as above using log-Poisson regression (appendix p 11). Weights were taken as the inverse of the modeled probability for participants-visits with resistance predictions, normalized to the total number of viremic PLHIV in each round. When estimating prevalence among all participants, people not living with HIV and non-viremic PLHIV were assigned a weight of 1.

Prevalence of viremic treatment-experienced resistance
The prevalence of NNRTI, NRTI, and PI resistance among viremic treatment-experienced PLHIV in rounds 17 and 18 was estimated using robust log-Poisson regression among participant-visits with available resistance prediction data to a given class using IPW to account for missing resistance predictions. The population prevalence of viremic treatment-experienced resistance in rounds 17 and 18 was calculated among all viremic treatment-experienced PLHIV, non-viremic treatment-experienced PLHIV, pre-treatment PLHIV, and participants not living with HIV with IPW to approximate the population of all RCCS participants.

Among participant-visits from viremic treatment-experienced PLHIV the probability of having NNRTI, NRTI, and PI resistance prediction was modeled based on the same predictors as above using log-Poisson regression (appendix p 14). Weights were taken as the inverse of the modeled probability for participants-visits with resistance predictions, normalized to the total number of viremic treatment-experienced PLHIV in each round. When estimating population-prevalence, people not living with HIV, pre-treatment PLHIV, and non-viremic treatment-experienced PLHIV were assigned a weight of 1.

Repeated measures
Correlation between repeated measures from the same participant were evaluated using geepack v.1.3.9.[2] We evaluated the impact of the assumed correlation structure on the estimated prevalence of resistance among viremic pre-treatment PLHIV, viremic treatment-experienced PLHIV, and all viremic PLHIV. Independent, exchangeable, AR1, and unstructured correlation structures were evaluated among viremic pre-treatment PLHIV in round 15 through 19 and independent and exchangeable correlation structures were evaluated among viremic treatment-experienced and all viremic PLHIV in round 17 and 18 (appendix pp 17-21). Due to the limited number of repeated measures, the QIC and CIC values were highly similar regardless of correlation structure. Further, the independent

correlation structure generated the widest confidence intervals. Consequently, independent correlation structure with a robust variance was assumed for all downstream analyses.

Identification of multi-class resistance
Among participant-visits from viremic PLHIV with genotype data for all NNRTIs, NRTIs, and PIs we tabulated the number of participants with all potential combinations of susceptibility/resistance for these three drug classes stratified by whether those participant-visits were pre-treatment or treatment-experienced. Further, in survey round 18 we estimated the prevalence of NNRTI single-class and multi-class resistance among all viremic PLHIV using robust log-Poisson regression with inverse probability weighting to account for the availability of resistance prediction to all NNRTI, NRTI, and PIs.

Weights were calculated by modeling the probability of having a resistance prediction to all NNRTIs, NRTIs, PIs, and INSTIs based on the same predictors as above among viremic pre-treatment and viremic treatment experienced round 18 participants using log-Poisson regression (appendix p 11). Weights were taken as the inverse of the modeled probability for participants-visits with resistance predictions, normalized to the total number of viremic pre-treatment and viremic treatment-experienced participants in each round.

Prevalence of viremic HIV and viremic resistance stratified by community
Robust log-Poisson regression was used to estimate the prevalence of viremic HIV among round 18 participants stratified by community type and individual community ID. Using data from round 18 participants who were not living with HIV, non-viremic PLHIV, and viremic PLHIV with available resistance we further estimated the prevalence of resistant viremia stratified by community type and community ID. When estimating the prevalence of viremic resistance IPW was used to account for the fact that some viremic PLHIV were missing resistance predictions.

Among participant-visits from viremic PLHIV, the probability of having NNRTI, NRTI, and PI resistance prediction was modeled based on the same predictors as above using log-Poisson regression. Weights were taken as the inverse of the modeled probability for participants-visits with resistance predictions, normalized to the total number of viremic PLHIV in each round. Participants not living with HIV non-viremic PLHIV were assigned a weight of 1.

Using the modeled prevalence of NNRTI, NRTI, and PI resistance among all viremic PLHIV we calculated the expected prevalence of resistant viremia among all participants, regardless of HIV status, as a product of the prevalence of resistance among viremic PLHIV and theoretical prevalence of viremia measures up to 16.5%.

Prevalence of amino acid substitutions associated with drug resistance
We first tabulated the number of participant visits contributed by viremic PLHIV between rounds 15 and 19 in which each individual drug resistance substitution was observed. Next, using pre-treatment viremic participant-visits in rounds 15 through 19 in which a resistance prediction was available for all drugs we estimated the prevalence of each individual amino acid substitution stratified by round using robust log-Poisson regression. IPW was used to adjust for missing resistance predictions among some pre-treatment viremic participant visits.

Using treatment-experienced viremic round 18 participants with resistance prediction to at least one drug we estimated the prevalence of each resistance substitution using robust log-Poisson regression. IPW was used to account for missing resistance predictions among some treatment-experienced viremic round 18 participants.

Weights were calculated by modeling the probability of having a resistance prediction to at least one drug based on the same predictors as above among viremic pre-treatment and viremic treatment experienced round 18 participants using log-Poisson regression (appendix pp 12-14). Weights were taken as the inverse of the modeled probability for participants-visits with resistance predictions, normalized to the total number of viremic pre-treatment and viremic treatment-experienced participants in each round.

The prevalence of the inT97A substitution among viremic pre-treatment PLHIV was further stratified by participant sex, age category, community type, and best-reference subtype in each round using robust log-Poisson regression with IPW as above. To stratify inT97A prevalence by subtype, best-reference subtypes observed in less than 300 participant-visits between round 15 and 19 were grouped into “other” categories based on whether they were clonal or recombinant subtypes. The stratified prevalence of inT97A was only estimated in the three most common

subtypes (subtypes A, D, and recombinant). We further tabulated the total number of participant-visits contributed by viremic PLHIV with and without the inT97A substitution stratified by best reference subtype, genome subtype, and polymerase subtype grouping all subtypes with less than three non-inT97A samples into "other." The distribution of $\log_{10}$ viral load measurements of all viremic participant-visits was calculated in 0.2 log10 copies/mL bins stratified by presence/absence of inT97A. The distribution of within-host frequency for inT97A and all other resistance substitutions was calculated in 5% bins and the distribution of support reading counts for inT97A and all other resistance substitutions was calculated in 1000 read bins.

Supplementary Table 2: Technology used to generate sequence data by survey round

	All	MiSeq/HiSeq	NovaSeq 6000 (veSEQ-HIV)
Overall	4263	2263 (53.08%)	2000 (46.92%)
Survey round			
15 (2012)	1822	1719 (94.35%)	103 (5.65%)
16 (2014)	692	532 (76.88%)	160 (23.12%)
17 (2015)	937	12 (1.28%)	925 (98.72%)
18 (2017)	631	0 (0%)	631 (100%)
19 (2019)	181	0 (0%)	181 (100%)

Supplementary Table 3: Drugs to which resistance was predicted

Generic name	Abbreviation	Class
doravirine	DOR	nnrti
efavirenz	EFV	nnrti
Etravirine	ETR	nnrti
nevirapine	NVP	nnrti
rilpivirine	RPV	nnrti
abacavir	ABC	nrti
zidovudine	AZT	nrti
stavudine	D4T	nrti
didanosine	DDI	nrti
emtricitabine	FTC	nrti
lamivudine	3TC	nrti
tenofovir	TDF	nrti
atazanavir	ATV	pi
darunavir	DRV	pi
fosamprenavir	FPV	pi
indinavir	IDV	pi
loprinavir	LPV	pi
nelfinavir	NFV	pi
saquinavir	SQV	pi
tipranavir	TPV	pi
bictegravir	BIC	insti
dolutegravir	DTG	insti
elvitegravir	EVG	insti
raltegravir	RAL	insti
cabotegravir	CAB	insti

NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = protease inhibitors.

Supplementary Table 4: Summary of outcome measures

Measure	Model outcome (Numerator)	Population (Denominator)	Pseudo-population (IPW population)	Model Predictors	Corresponding figure
population prevalence of PLHIV in RCCS	HIV seropositive	all RCCS participants	all RCCS participants	survey round (R15/R16/R17/R18/R19)	Fig 1A
population prevalence of viremic PLHIV in RCCS	viremic and HIV seropositive			survey round (R16/R17/R18/R19)	Fig 1A
population prevalence of viremic pre-treatment PLHIV in RCCS	viremic pre-treatment, and HIV seropositive			survey round (R15/R16/R17/R18/R19)	Fig 1A
population prevalence of viremic treatment-experienced PLHIV in RCCS	viremic treatment-experienced, and HIV seropositive			survey round (R16/R17/R18/R19)	Fig 1A
population prevalence of viremia in R18	viremic and HIV seropositive			community type (agrarian/fishing/trading)	Fig 4C,D
population prevalence of viremic pre-treatment NNRTI/NRTI/PI resistance	viremic pre-treatment NNRTI/NRTI/PI resistance	viremic pre-treatment PLHIV with resistance prediction to all NNRTIs/NRTIs/Pis; non-viremic pre-treatment PLHIV; treatment-experienced PLHIV; participants not living with HIV		survey round (R15/R16/R17/R18/R19)	Fig 1B
population prevalence of viremic treatment-experienced NNRTI/NRTI/PI resistance	viremic treatment-experienced NNRTI/NRTI/PI resistance	viremic treatment-experienced PLHIV with resistance prediction to all NNRTIs/NRTIs/Pis; non-viremic treatment-experienced PLHIV; pre-treatment PLHIV; participants not living with HIV		survey round (R17/R18)	Fig 1C
population prevalence of viremic NNRTI/NRTI/PI resistance	viremic NNRTI/NRTI/PI resistance	viremic PLHIV with resistance prediction to all NNRTIs, NRTIs, and PIs; non-viremic PLHIV; participants not living with HIV		survey round (R17/R18)	
population prevalence of viremic NNRTI/NRTI/PI resistance in R18				community type (agrarian/fishing/trading)	Fig 4A,B
population prevalence of viremic NNRTI/NRTI/PI resistance in R18				community ID	Fig 4A,B
population prevalence of multi-class NNRTI resistance in R18	NNTI resistance and resistance to NRTIs and/or INSTIs				Fig 2C
population prevalence of NNRTI mono-resistance in R18	NNTI resistance and not resistance to NRTIs and INSTIs				
prevalence of NNRTI/NRTI/PI resistance among viremic pre-treatment PLHIV	viremic NNRTI/NRTI/PI resistance	viremic pre-treatment PLHIV with resistance prediction to all NNRTIs/NRTIs/PIs	viremic pre-treatment PLHIV	survey round (R15/R16/R17/R18/R19)	Fig 1D
prevalence of NNRTI resistance among viremic pre-treatment PLHIV	NNRTI resistance	viremic pre-treatment PLHIV with resistance prediction to all NNRTIs		survey round × community type (agrarian/fishing/trading)	Fig S4
prevalence of NNRTI resistance among viremic pre-treatment PLHIV				survey round × sex (M/F)	Fig S4
prevalence of NNRTI resistance among viremic pre-treatment PLHIV				survey round × age category ((14,24]/(24,34]/(34,50])	Fig S4

prevalence of resistance substitutions* among viremic pre-treatment PLHIV	substitution presence	viremic pre-treatment PLHIV with resistance prediction to all NNRTIs, NRTIs, PIs, and INSTIs		survey round (R15/R16/R17/R18/R19)	Fig 4B,C,E
prevalence of inT97A among viremic pre-treatment PLHIV	inT97A presence			survey round × community type (agrarian/fishing/trading)	Fig S6
prevalence of inT97A among viremic pre-treatment PLHIV				survey round × sex (M/F)	Fig S6
prevalence of inT97A among viremic pre-treatment PLHIV				survey round × age category ((14,24]/(24,34]/(34,50])	Fig S6
prevalence of inT97A among viremic pre-treatment PLHIV				survey round × subtype (A1, D, other (recombinant))*	Fig S6
probability of NNRTI resistance among sero-incident, pre-treatment PLHIV	NNRTI resistance at first RCCS visit with available NNRTI resistance prediction	sero-incident PLHIV in the RCCS with last negative and first positive in conseucitve RCCS rounds and available pre-treatment resistance prediction to all NNRTIs	sero-incident PLHIV in the RCCS with last negative and first positive in conseucitve RCCS rounds	survey round (R15/R16/R17/R18/R19)	
prevalence of NNRTI/NRI/PI resistance among viremic treatment-experienced PLHIV	NNRTI/NRTI/PI resistance	viremic treatment-experiencedPLHIV with resistance prediction to all NNRTIs/NRTIs/Pis	viremic treatment-experienced PLHIV	Survey round (R17/R18)	Fig 1E
prevalence of resistance substitions* among viremic treatment-experienced PLHIV in R18	substitution presence	viremic treatment-experienced PLHIV with resistance prediction to all NNRTis, NRTIs, Pis, and INSTIs			Fig 4D,E

Measure refers to how a given analysis is referred to in text, figures, and tables. The model outcome is the participant-visit level characteristic which is being modeled in each analysis. The population is the data among which a given outcome is modeled. Pseudo-population is the population which is being approximated using IPW. Model predictors are the variables included in each model. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = protease inhibitors. INSTI = integrase strand transfer inhibitors. IPW = inverse probability weighting.

Supplementary Table 5: Inverse probability weighting model for the probability that genotyping data is available among viremic PLHIV

Variable	NNRTI		NRTI		PI		NNRTI, NRTIs, and PIs	
	Coefficient (95% CI)	*p*-value	Coefficient (95% CI)	*p*-value	Coefficient (95% CI)	*p*-value	Coefficient (95% CI)	*p*-value
(Intercept)	0.45 (0.37, 0.54)	<0.0001	0.3 (0.24, 0.37)	<0.0001	0.42 (0.35, 0.51)	<0.0001	0.26 (0.2, 0.33)	<0.0001
Survey round								
round 17	ref	ref	ref	ref	ref	ref	ref	ref
round 18	1.76 (1.34, 2.31)	<0.0001	2.19 (1.61, 2.97)	<0.0001	1.85 (1.41, 2.44)	<0.0001	2.54 (1.85, 3.49)	<0.0001
Age category								
(14, 24] years old	ref	ref	ref	ref	ref	ref	ref	ref
(24,34] years old	0.96 (0.9, 1.03)	0.29	0.91 (0.84, 0.98)	0.02	0.95 (0.88, 1.01)	0.11	0.93 (0.85, 1.01)	0.08
(34,50] years old	0.96 (0.89, 1.04)	0.31	0.94 (0.87, 1.03)	0.17	0.98 (0.91, 1.05)	0.51	0.97 (0.88, 1.06)	0.47
(24,34] years old × round 18	1.01 (0.91, 1.11)	0.88	1.04 (0.93, 1.16)	0.49	1.01 (0.92, 1.12)	0.78	1.02 (0.91, 1.14)	0.77
(34,50] years old × round 18	0.96 (0.86, 1.08)	0.51	0.95 (0.84, 1.07)	0.37	0.94 (0.84, 1.06)	0.32	0.92 (0.81, 1.04)	0.18
Community type								
Agrarian community	ref	ref	ref	ref	ref	ref	ref	ref
Fishing community	1.02 (0.96, 1.09)	0.46	1.01 (0.94, 1.09)	0.76	1.01 (0.95, 1.08)	0.70	0.99 (0.92, 1.07)	0.89
Trading community	1.01 (0.94, 1.09)	0.78	1.06 (0.97, 1.15)	0.19	1.03 (0.96, 1.12)	0.37	1.04 (0.96, 1.14)	0.34
Fishing community × round 18	1 (0.91, 1.1)	0.93	1.04 (0.94, 1.16)	0.45	1.03 (0.93, 1.13)	0.60	1.06 (0.95, 1.18)	0.29
Trading community × round 18	1.02 (0.91, 1.14)	0.77	0.98 (0.87, 1.11)	0.77	1.01 (0.9, 1.13)	0.91	0.99 (0.87, 1.12)	0.84
Sex								
Female	ref	ref	ref	ref	ref	ref	ref	ref
Male	0.97 (0.92, 1.03)	0.34	0.98 (0.91, 1.04)	0.46	0.95 (0.9, 1.01)	0.08	0.97 (0.9, 1.03)	0.33
Male × round 18	1 (0.92, 1.09)	0.94	0.99 (0.9, 1.09)	0.87	1.02 (0.94, 1.11)	0.60	1 (0.91, 1.1)	0.97
Viral load (log10 copies/mL)	1.17 (1.12, 1.22)	<0.0001	1.28 (1.22, 1.34)	<0.0001	1.19 (1.14, 1.24)	<0.0001	1.32 (1.25, 1.39)	<0.0001
Viral load (log10 copies/mL) × round 18	0.88 (0.83, 0.93)	<0.0001	0.84 (0.78, 0.9)	<0.0001	0.86 (0.81, 0.92)	<0.0001	0.81 (0.76, 0.87)	<0.0001

Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. *p*-values that coefficients are different from 0 at the $\alpha = 0.05$ level were calculated using the Wald method. Round 17 participant-visits from women aged 14 to 24 in Agrarian communities with a $\log_{10}$ viral load of 0 was used as the reference category. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = protease inhibitors.

Supplementary Table 6: Inverse probability weighting model for the probability that genotyping data is available among viremic pre-treatment PLHIV

Variable	NNRTI		NRTI		PI		All drug classes	
	Coefficient (95% CI)	*p*-value	Coefficient (95% CI)	*p*-value	Coefficient (95% CI)	*p*-value	Coefficient (95% CI)	*p*-value
(Intercept)	0.04 (0.02, 0.06)	<0.0001	0.03 (0.02, 0.05)	<0.0001	0.13 (0.1, 0.18)	<0.0001	0 (0, 0.01)	<0.0001
Survey round								
round 15	ref	ref	ref	ref	ref	ref	ref	ref
round 16	0.58 (0.3, 1.11)	0.1	0.54 (0.28, 1.04)	0.065	0.38 (0.22, 0.64)	0.00032	1.66 (0.62, 4.46)	0.32
round 17	12.76 (7.9, 20.61)	<0.0001	9.75 (6, 15.85)	<0.0001	3.58 (2.48, 5.16)	<0.0001	115.74 (52.21, 256.57)	<0.0001
round 18	21.4 (13.32, 34.37)	<0.0001	19.25 (11.91, 31.11)	<0.0001	6.02 (4.2, 8.62)	<0.0001	263.13 (119.51, 579.32)	<0.0001
round 19	32.7 (11.78, 90.78)	<0.0001	31.32 (11.19, 87.67)	<0.0001	11.61 (4.48, 30.08)	<0.0001	428.09 (128.3, 1428.36)	<0.0001
Age category								
(14, 24] years old	ref	ref	ref	ref	ref	ref	ref	ref
(24,34] years old	0.98 (0.87, 1.11)	0.77	0.99 (0.89, 1.12)	0.93	1 (0.92, 1.08)	0.94	0.92 (0.75, 1.12)	0.40
(34,50] years old	0.92 (0.81, 1.06)	0.24	0.96 (0.84, 1.09)	0.50	0.98 (0.89, 1.07)	0.63	0.89 (0.71, 1.11)	0.29
(24,34] years old × round 16	0.76 (0.61, 0.94)	0.01	0.77 (0.62, 0.95)	0.02	0.75 (0.62, 0.9)	0.00	0.82 (0.59, 1.14)	0.24
(24,34] years old × round 17	1.04 (0.9, 1.19)	0.61	0.95 (0.82, 1.1)	0.51	1 (0.9, 1.12)	0.95	1.08 (0.87, 1.35)	0.50
(24,34] years old × round 18	0.95 (0.83, 1.1)	0.51	0.91 (0.79, 1.05)	0.19	0.93 (0.83, 1.04)	0.22	0.99 (0.79, 1.22)	0.90
(24,34] years old × round 19	0.96 (0.67, 1.39)	0.83	0.99 (0.68, 1.44)	0.94	0.93 (0.66, 1.31)	0.67	1.07 (0.71, 1.62)	0.75
(34,50] years old × round 16	0.73 (0.57, 0.93)	0.01	0.76 (0.6, 0.97)	0.03	0.71 (0.58, 0.87)	0.00	0.82 (0.58, 1.18)	0.29
(34,50] years old × round 17	1.06 (0.91, 1.25)	0.44	1 (0.85, 1.17)	1.00	1.03 (0.91, 1.16)	0.65	1.12 (0.88, 1.43)	0.36
(34,50] years old × round 18	1.03 (0.88, 1.21)	0.74	0.94 (0.8, 1.11)	0.46	0.96 (0.84, 1.09)	0.50	1.02 (0.8, 1.3)	0.90
(34,50] years old × round 19	1.04 (0.7, 1.55)	0.85	1.02 (0.67, 1.54)	0.93	0.97 (0.66, 1.4)	0.86	1.1 (0.7, 1.73)	0.68
Community type								
Agrarian community	ref	ref	ref	ref	ref	ref	ref	ref
Fishing community	0.84 (0.71, 0.98)	0.03	0.89 (0.76, 1.05)	0.17	1.05 (0.94, 1.19)	0.38	0.9 (0.66, 1.23)	0.51
Trading community	0.95 (0.84, 1.09)	0.49	0.97 (0.85, 1.1)	0.63	0.99 (0.9, 1.1)	0.91	1.01 (0.81, 1.25)	0.96
Fishing community × round 16	0.82 (0.65, 1.04)	0.11	0.86 (0.68, 1.1)	0.23	0.71 (0.58, 0.86)	0.00	0.6 (0.4, 0.88)	0.01
Fishing community × round 17	1.24 (1.04, 1.47)	0.02	1.17 (0.98, 1.4)	0.09	0.98 (0.85, 1.12)	0.76	1.14 (0.83, 1.58)	0.41
Fishing community × round 18	1.18 (0.99, 1.42)	0.06	1.14 (0.95, 1.37)	0.15	0.95 (0.83, 1.1)	0.52	1.13 (0.82, 1.56)	0.45
Fishing community × round 19	1.55 (1.09, 2.2)	0.02	1.4 (0.98, 2)	0.07	1.3 (0.94, 1.8)	0.12	1.38 (0.89, 2.16)	0.15

Trading community × round 16	1.04 (0.82, 1.31)	0.75	1.1 (0.87, 1.39)	0.44	1.12 (0.92, 1.36)	0.26	0.98 (0.71, 1.37)	0.92
Trading community × round 17	1.04 (0.89, 1.22)	0.62	1.09 (0.93, 1.28)	0.30	1.05 (0.92, 1.19)	0.48	1.03 (0.81, 1.31)	0.81
Trading community × round 18	1.06 (0.91, 1.24)	0.45	1.04 (0.89, 1.23)	0.60	1.03 (0.9, 1.17)	0.70	1.01 (0.79, 1.28)	0.96
Trading community × round 19	0.97 (0.66, 1.43)	0.87	0.96 (0.65, 1.41)	0.84	0.92 (0.63, 1.35)	0.68	0.93 (0.6, 1.42)	0.72
Sex								
Female	ref	ref	ref	ref	ref	ref	ref	ref
Male	1.11 (1.01, 1.21)	0.03	1.12 (1.02, 1.22)	0.01	1.1 (1.03, 1.17)	0.00	1.1 (0.94, 1.28)	0.25
Male × round 16	0.76 (0.64, 0.92)	0.00	0.77 (0.64, 0.92)	0.00	0.82 (0.7, 0.95)	0.01	0.79 (0.61, 1.02)	0.08
Male × round 17	0.87 (0.78, 0.96)	0.01	0.86 (0.77, 0.97)	0.01	0.86 (0.79, 0.94)	0.00	0.87 (0.73, 1.03)	0.11
Male × round 18	0.89 (0.8, 1)	0.05	0.88 (0.78, 0.99)	0.03	0.89 (0.81, 0.98)	0.02	0.9 (0.75, 1.07)	0.22
Male × round 19	0.81 (0.6, 1.08)	0.15	0.8 (0.59, 1.07)	0.13	0.83 (0.63, 1.09)	0.18	0.81 (0.59, 1.12)	0.21
Viral load								
Missing	11.64 (7.66, 17.7)	<0.0001	12.95 (8.63, 19.46)	<0.0001	4.16 (3.12, 5.55)	<0.0001	98.46 (47.27, 205.08)	<0.0001
Viral load (log10 copies/mL)	1.87 (1.72, 2.04)	<0.0001	1.9 (1.75, 2.07)	<0.0001	1.46 (1.37, 1.54)	<0.0001	2.89 (2.5, 3.33)	<0.0001
Viral load (log10 copies/mL) × round 16	1.13 (0.99, 1.3)	0.076	1.13 (0.99, 1.29)	0.079	1.22 (1.09, 1.37)	0.00043	0.97 (0.8, 1.19)	0.8
Viral load (log10 copies/mL) × round 17	0.62 (0.56, 0.68)	<0.0001	0.66 (0.6, 0.73)	<0.0001	0.79 (0.74, 0.85)	<0.0001	0.44 (0.38, 0.52)	<0.0001
Viral load (log10 copies/mL) × round 18	0.55 (0.5, 0.61)	<0.0001	0.57 (0.52, 0.63)	<0.0001	0.71 (0.66, 0.77)	<0.0001	0.38 (0.32, 0.44)	<0.0001
Viral load (log10 copies/mL) × round 19	0.42 (0.34, 0.52)	<0.0001	0.43 (0.34, 0.53)	<0.0001	0.51 (0.42, 0.63)	<0.0001	0.28 (0.22, 0.36)	<0.0001

Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. *p*-values that coefficients are different from 0 at the $\alpha = 0.05$ level were calculated using the Wald method. Round 15 participant-visits from women aged 14 to 24 in Agrarian communities with a $\log_{10}$ viral load of 0 was used as the reference category. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = protease inhibitors. INSTI = integrase strand transfer inhibitors. All drug classes = genotype data for all INSTIs, NNRTIs, NRTIs, and PIs.

Supplementary Table 7: Inverse probability weighting model for the probability that genotyping data is available among viremic treatment-experienced PLHIV

Variable	NNRTI		NRTI		PI		All drug classes	
	Coefficient (95% CI)	*p*-value	Coefficient (95% CI)	*p*-value	Coefficient (95% CI)	*p*-value	Coefficient (95% CI)	*p*-value
(Intercept)	0.36 (0.22, 0.61)	0.00015	0.23 (0.13, 0.38)	<0.0001	0.29 (0.18, 0.47)	<0.0001	0.17 (0.09, 0.33)	<0.0001
Survey round								
round 17	ref	ref	ref	ref	ref	ref	ref	ref
round 18	1.93 (0.96, 3.89)	0.067	2.79 (1.36, 5.72)	0.005	2.43 (1.25, 4.75)	0.0091	3.62 (1.62, 8.05)	0.0016
Age category								
(14, 24] years old	ref	ref	ref	ref	ref	ref	ref	ref
(24,34] years old	0.75 (0.62, 0.91)	0.00	0.75 (0.61, 0.92)	0.01	0.73 (0.61, 0.88)	0.00	0.72 (0.56, 0.93)	0.01
(34,50] years old	0.86 (0.72, 1.04)	0.13	0.88 (0.73, 1.08)	0.22	0.87 (0.73, 1.04)	0.12	0.9 (0.71, 1.16)	0.42
(24,34] years old × round 18	1.49 (1.13, 1.96)	0.00	1.46 (1.09, 1.95)	0.01	1.52 (1.16, 1.98)	0.00	1.52 (1.1, 2.09)	0.01
(34,50] years old × round 18	1.1 (0.82, 1.47)	0.54	1.07 (0.79, 1.44)	0.66	1.12 (0.84, 1.49)	0.42	1.03 (0.74, 1.44)	0.85
Community type								
Agrarian community	ref	ref	ref	ref	ref	ref	ref	ref
Fishing community	0.98 (0.84, 1.16)	0.83	0.89 (0.75, 1.07)	0.21	0.95 (0.81, 1.11)	0.51	0.86 (0.72, 1.04)	0.11
Trading community	1.06 (0.87, 1.29)	0.59	1.05 (0.86, 1.28)	0.66	0.99 (0.81, 1.2)	0.89	0.96 (0.75, 1.22)	0.72
Fishing community × round 18	1.17 (0.93, 1.47)	0.19	1.3 (1.01, 1.66)	0.04	1.21 (0.96, 1.52)	0.11	1.34 (1.05, 1.73)	0.02
Trading community × round 18	1.02 (0.76, 1.37)	0.87	1.09 (0.81, 1.46)	0.56	1.13 (0.85, 1.51)	0.41	1.16 (0.83, 1.61)	0.38
Sex								
Female	ref	ref	ref	ref	ref	ref	ref	ref
Male	1.01 (0.86, 1.19)	0.91	1 (0.84, 1.19)	0.98	0.95 (0.8, 1.12)	0.53	0.99 (0.82, 1.19)	0.92
Male × round 18	0.92 (0.75, 1.14)	0.44	0.93 (0.74, 1.16)	0.51	0.99 (0.8, 1.23)	0.96	0.92 (0.73, 1.16)	0.48
Viral load (log10 copies/mL)	1.26 (1.1, 1.44)	0.00078	1.41 (1.24, 1.6)	<0.0001	1.34 (1.2, 1.51)	<0.0001	1.49 (1.27, 1.75)	<0.0001
Viral load (log10 copies/mL) × round 18	0.81 (0.68, 0.95)	0.01	0.74 (0.63, 0.86)	0.00017	0.75 (0.65, 0.87)	0.00021	0.7 (0.58, 0.84)	0.00016

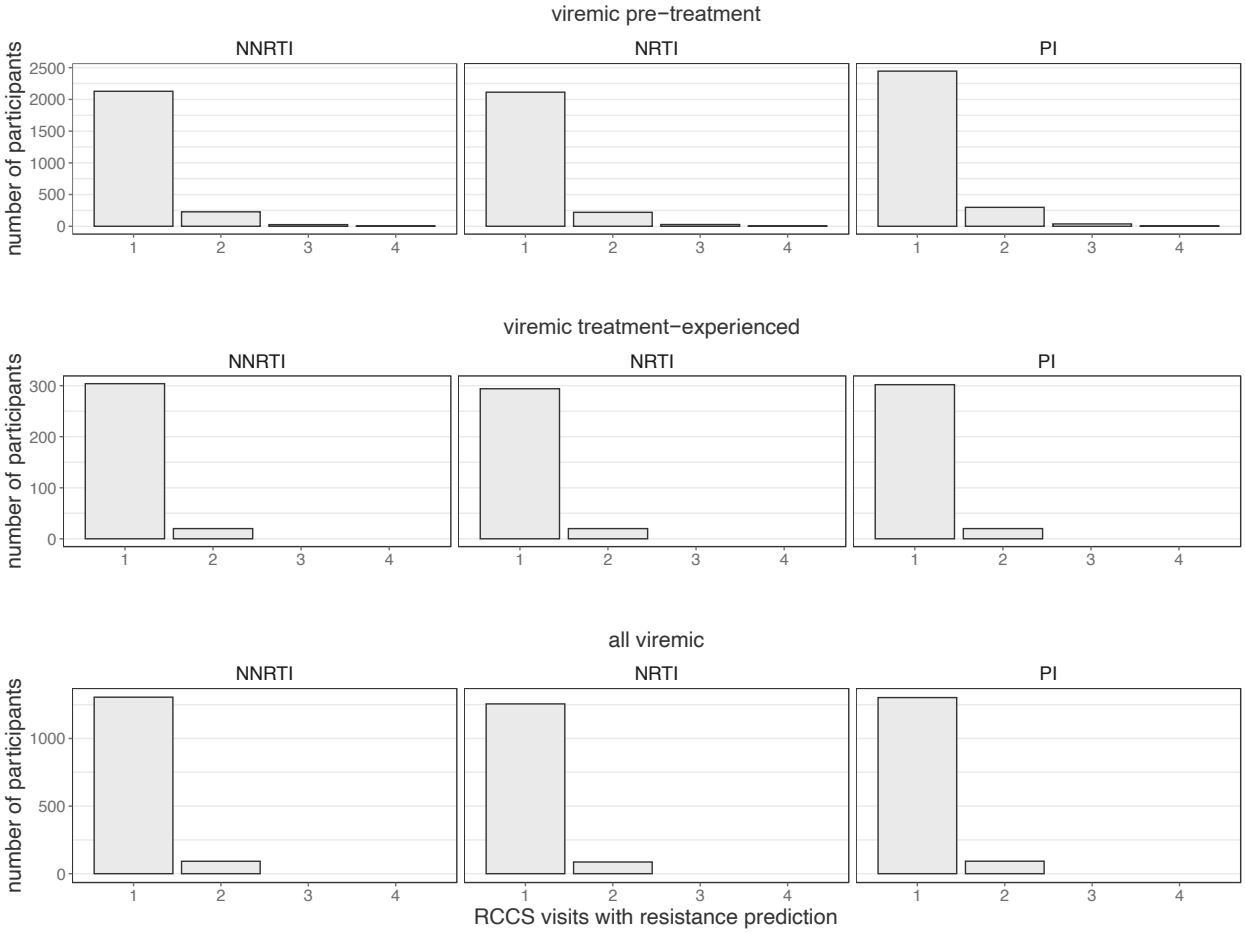
Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. *p*-values that coefficients are different from 0 at the $\alpha = 0.05$ level were calculated using the Wald method. Round 17 participant-visits from women aged 14 to 24 in Agrarian communities with a $\log_{10}$ viral load of 0 was used as the reference category. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = protease inhibitors. All drug classes = genotype data for all INSTIs, NNRTIs, NRTIs, and PIs.

Supplementary Table 8: Inverse probability weighting model for the probability that genotyping data is available among seroincident viremic pre-treatment PLHIV

	NNRTI	
Variable	**Coefficient (95% CI)**	***p*-value**
(Intercept)	0.02 (0, 0.18)	0.0008
Survey round		
round 15	ref	ref
round 16	4.17 (0.27, 63.81)	0.3
round 17	10.47 (0.82, 133.81)	0.071
round 18	38.71 (3.31, 452.56)	0.0036
round 19	149.98 (6.49, 3466.65)	0.0018
Age category		
(14, 24] years old	ref	ref
(24,34] years old	1.68 (0.97, 2.91)	0.07
(34,50] years old	1.48 (0.83, 2.62)	0.18
(24,34] years old × round 16	0.59 (0.27, 1.29)	0.19
(24,34] years old × round 17	0.64 (0.35, 1.19)	0.16
(24,34] years old × round 18	0.53 (0.3, 0.94)	0.03
(24,34] years old × round 19	0.9 (0.38, 2.1)	0.80
(34,50] years old × round 16	0.72 (0.31, 1.64)	0.43
(34,50] years old × round 17	0.74 (0.39, 1.4)	0.35
(34,50] years old × round 18	0.6 (0.33, 1.08)	0.09
(34,50] years old × round 19	0.8 (0.3, 2.08)	0.64
Sex		
Female	ref	ref
Male	0.69 (0.47, 1.02)	0.07
Male × round 16	1.44 (0.81, 2.58)	0.22
Male × round 17	1.22 (0.78, 1.92)	0.38
Male × round 18	1.3 (0.84, 2.01)	0.24
Male × round 19	1.29 (0.66, 2.5)	0.45
Viral load		
Missing	28.45 (2.61, 310.31)	0.006
Viral load (log10 copies/mL)	2.15 (1.31, 3.52)	0.0023
Viral load (log10 copies/mL) × round 16	0.74 (0.43, 1.29)	0.29
Viral load (log10 copies/mL) × round 17	0.66 (0.39, 1.11)	0.12
Viral load (log10 copies/mL) × round 18	0.52 (0.31, 0.87)	0.013
Viral load (log10 copies/mL) × round 19	0.32 (0.16, 0.62)	0.00078

Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. *p*-values that coefficients are different from 0 at the $\alpha = 0.05$ level were calculated using the Wald method. Women aged 14 to 24 in Agrarian communities with a $\log_{10}$ viral load of 0 whose earliest drug resistance prediction was from round 15 was used as the reference category. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors.

Supplementary Figure 2: Number of participant-visits with resistant predictions for each participant



Pre-treatment data include survey rounds 15 through 19. Viremic treatment-experienced and all viremic data include survey rounds 17 and 18. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = integrase inhibitors.

Supplementary Table 9: Prevalence of resistance among viremic PLHIV correlation structure model comparison

			NNRTI			NRTI			PI		
	Correlation Structure	**Params**	**Quasi-likelihood**	**QIC**	**CIC**	**Quasi-likelihood**	**QIC**	**CIC**	**Quasi-likelihood**	**QIC**	**CIC**
Viremic pre-treatment PLHIV	independence	5	-660.979	1344.565	11.304	-244.683	510.772	10.703	-280.846	582.07	10.189
	exchangeable	5	-661.098	1341.888	9.846	-244.742	510.785	10.65	-280.93	581.708	9.924
	ar1	5	-661.066	1341.753	9.811	-244.758	511.079	10.782	-280.906	580.967	9.578
	unstructured	5	-661.177	1342.27	9.958	-244.796	511.775	11.091	-280.901	580.575	9.387
Viremic treatment-experienced PLHIV	independence	2	-292.229	589.548	2.545	-260.297	525.872	2.639	-41.588	88.141	2.482
	exchangeable	2	-293.07	589.201	1.53	-260.943	526.88	2.497	-41.587	88.108	2.467
Viremic PLHIV	independence	2	-723.392	1451.46	2.338	-512.421	1029.641	2.399	-135.047	274.911	2.408
	exchangeable	2	-724.675	1453.185	1.917	-513.354	1032.279	2.786	-135.06	274.231	2.055

Estimates were generated using robust log-Poisson regression with generalized estimating equation. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = integrase inhibitors. QIC = quasi information criterion, CIC = correlation information criterion.

Supplementary Table 10: Prevalence of resistance among viremic pre-treatment PLHIV by correlation structure

		Independence		Exchangeable		AR1		Unstructured	
	Coefficient	Estimate	Std. Err	Estimate	Std. Err	Estimate	Std. Err	Estimate	Std. Err
NNRTI	(Intercept)	-3.001	0.149	-2.978	0.141	-2.979	0.141	-2.975	0.14
	round 16	0.39	0.243	0.379	0.219	0.373	0.218	0.391	0.216
	round 17	0.363	0.2	0.368	0.177	0.366	0.177	0.353	0.175
	round 18	0.675	0.207	0.642	0.197	0.666	0.195	0.603	0.202
	round 19	0.568	0.34	0.572	0.317	0.569	0.319	0.557	0.319
NRTI	(Intercept)	-4.088	0.227	-4.098	0.229	-4.101	0.23	-4.103	0.23
	round 16	-0.123	0.503	-0.123	0.515	-0.121	0.518	-0.122	0.527
	round 17	0.1	0.35	0.065	0.353	0.062	0.357	0.042	0.364
	round 18	0.051	0.418	0.108	0.397	0.106	0.4	0.098	0.405
	round 19	-0.018	0.745	0.019	0.717	0.016	0.723	0.012	0.728
PI	(Intercept)	-3.862	0.189	-3.866	0.189	-3.869	0.19	-3.871	0.19
	round 16	-0.101	0.423	-0.068	0.419	-0.071	0.422	-0.07	0.424
	round 17	-0.467	0.383	-0.414	0.358	-0.411	0.358	-0.421	0.36
	round 18	-0.039	0.373	-0.075	0.374	-0.031	0.359	-0.03	0.358
	round 19	-0.935	1.014	-0.956	1.034	-0.896	0.975	-0.838	0.924

Estimates were generated using robust log-Poisson regression with generalized estimating equation. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = integrase inhibitors.

Supplementary Table 11: Prevalence of resistance among viremic treatment-experienced PLHIV by correlation structure

		Independence		Exchangeable	
	Coefficient	Estimate	Std. Err	Estimate	Std. Err
NNRTI	(Intercept)	-0.645	0.076	-0.746	0.061
	round16	-0.096	0.103	0.049	0.026
NRTI	(Intercept)	-0.767	0.089	-0.848	0.095
	round16	-0.24	0.127	-0.1	0.117
PI	(Intercept)	-3.532	0.451	-3.529	0.449
	round16	-0.292	0.668	-0.292	0.664

Estimates were generated using robust log-Poisson regression with generalized estimating equation. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = integrase inhibitors.

Supplementary Table 12: Prevalence of resistance among viremic PLHIV by correlation structure

		Independence		Exchangeable	
	Coefficient	Estimate	Std. Err	Estimate	Std. Err
NNRTI	(Intercept)	-1.858	0.079	-1.767	0.068
	round16	0.288	0.105	0.079	0.083
NRTI	(Intercept)	-2.286	0.103	-2.396	0.132
	round16	0.153	0.141	0.25	0.147
PI	(Intercept)	-4.107	0.268	-4.078	0.249
	round16	0.242	0.371	0.221	0.313

Estimates were generated using robust log-Poisson regression with generalized estimating equation. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = integrase inhibitors.

Supplementary Table 13: Prevalence of PLHIV, viremic PLHIV, viremic pre-treatment PLHIV, and viremic treatment-experienced PLHIV

	PLHIV			Viremic PLHIV		
Survey round	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**
15 (August 2011 - May 2013)	20.37 (19.78, 20.98)	ref	ref	..	..	..
16 (July 2013 - January 2015)	18.84 (18.27, 19.42)	0.925 (0.886, 0.965)	0.00028	8.14 (7.75, 8.55)	ref	ref
17 (February 2015 - September 2016)	18.71 (18.16, 19.26)	0.918 (0.881, 0.957)	<0.0001	5.3 (5, 5.63)	0.652 (0.603, 0.704)	<0.0001
18 (October 2016 - May 2018)	18.36 (17.83, 18.91)	0.901 (0.865, 0.94)	<0.0001	3.67 (3.42, 3.95)	0.452 (0.414, 0.492)	<0.0001
19 (June 2018 - November 2020)	17.2 (16.68, 17.74)	0.844 (0.809, 0.881)	<0.0001	2.34 (2.14, 2.57)	0.288 (0.26, 0.319)	<0.0001

	Viremic pre-treatment PLHIV			Viremic treatment-experienced PLHIV		
Survey round	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**
15 (August 2011 - May 2013)	13.08 (12.59, 13.6)	ref	ref	..	..	..
16 (July 2013 - January 2015)	7.33 (6.96, 7.72)	0.56 (0.525, 0.598)	<0.0001	0.81 (0.69, 0.95)	ref	ref
17 (February 2015 - September 2016)	4.25 (3.97, 4.54)	0.325 (0.301, 0.351)	<0.0001	1.05 (0.92, 1.21)	1.301 (1.053, 1.607)	0.015
18 (October 2016 - May 2018)	2.53 (2.32, 2.76)	0.194 (0.176, 0.213)	<0.0001	1.14 (1, 1.3)	1.407 (1.144, 1.73)	0.0012
19 (June 2018 - November 2020)	1.49 (1.33, 1.67)	0.114 (0.101, 0.129)	<0.0001	0.85 (0.73, 0.99)	1.052 (0.843, 1.314)	0.65

Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. *p*-values that coefficients are different from 0 at the $\alpha = 0.05$ level were calculated using the Wald method. PLHIV = people living with HIV.

Supplementary Table 14: Number of viremic samples at participant visits that had successful genotyping for all INSTIs

	Viremic PLWHIV			Viremic pre-treatment PLWHIV			Viremic treatment-experienced PLWHIV		
	Participant-visits	**Genotype data for all INSTIs (%)**	***p*-value**	**Participant-visits**	**Genotype data for all INSTIs (%)**	***p*-value**	**Participant-visits**	**Genotype data for all INSTIs (%)**	***p*-value**
Overall	5953	2593 (43.56%)		5177	2183 (42.17%)		776	410 (52.84%)	
Age Category			0.49			0.42			0.73
(14,24]	1146	528 (46.07%)		1032	464 (44.96%)		114	64 (56.14%)	
(24,34]	2931	1255 (42.82%)		2573	1061 (41.24%)		358	194 (54.19%)	
(34,50]	1876	810 (43.18%)		1572	658 (41.86%)		304	152 (50%)	
Community type			0.42			0.045			0.17
Agrarian	1794	803 (44.76%)		1590	709 (44.59%)		204	94 (46.08%)	
Fishing	2861	1206 (42.15%)		2455	966 (39.35%)		406	240 (59.11%)	
Trading	1298	584 (44.99%)		1132	508 (44.88%)		166	76 (45.78%)	
Sex			0.095			0.06			0.82
F	3188	1337 (41.94%)		2744	1104 (40.23%)		444	233 (52.48%)	
M	2765	1256 (45.42%)		2433	1079 (44.35%)		332	177 (53.31%)	
Survey round			<1e-4			<1e-4			<1e-4
15 (2012)	2281	551 (24.16%)		2246	551 (24.53%)		35	0 (0%)	
16 (2014)	1465	335 (22.87%)		1319	333 (25.25%)		146	2 (1.37%)	
17 (2015)	1026	909 (88.6%)		822	735 (89.42%)		204	174 (85.29%)	
18 (2017)	728	628 (86.26%)		502	442 (88.05%)		226	186 (82.3%)	
19 (2019)	453	170 (37.53%)		288	122 (42.36%)		165	48 (29.09%)	
Viral load (log10 copies/mL)			<1e-4			<1e-4			0.8
(3,4]	2101	816 (38.84%)		1748	628 (35.93%)		353	188 (53.26%)	
(4,5]	2231	1153 (51.68%)		1898	977 (51.48%)		333	176 (52.85%)	
(5,Inf]	598	362 (60.54%)		508	316 (62.2%)		90	46 (51.11%)	
missing	1023	262 (25.61%)		1023	262 (25.61%)				

Percentages show the percentage of all samples within a category that have successful genotyping for all INSTIs. *p*-values generated using the χ^2 test. PLWHIV=people living with HIV. INSTI = integrase strand transfer inhibitors.

Supplementary Table 15: Number of viremic samples at participant visits that had successful genotyping for all NNRTIs

	Viremic PLWHIV			Viremic pre-treatment PLWHIV			Viremic treatment-experienced PLWHIV		
	Participant-visits	Genotype data for all NNRTIs (%)	*p*-value	Participant-visits	Genotype data for all NNRTIs (%)	*p*-value	Participant-visits	Genotype data for all NNRTIs (%)	*p*-value
Overall	5953	3080 (51.74%)		5177	2686 (51.88%)		776	394 (50.77%)	
Age Category			0.27			0.36			0.73
(14,24]	1146	630 (54.97%)		1032	567 (54.94%)		114	63 (55.26%)	
(24,34]	2931	1520 (51.86%)		2573	1335 (51.88%)		358	185 (51.68%)	
(34,50]	1876	930 (49.57%)		1572	784 (49.87%)		304	146 (48.03%)	
Community type			0.98			0.79			0.17
Agrarian	1794	932 (51.95%)		1590	842 (52.96%)		204	90 (44.12%)	
Fishing	2861	1482 (51.8%)		2455	1253 (51.04%)		406	229 (56.4%)	
Trading	1298	666 (51.31%)		1132	591 (52.21%)		166	75 (45.18%)	
Sex			0.086			0.085			0.82
F	3188	1590 (49.87%)		2744	1368 (49.85%)		444	222 (50%)	
M	2765	1490 (53.89%)		2433	1318 (54.17%)		332	172 (51.81%)	
Survey round			<1e-4			<1e-4			<1e-4
15 (2012)	2281	1001 (43.88%)		2246	1000 (44.52%)		35	1 (2.86%)	
16 (2014)	1465	425 (29.01%)		1319	422 (31.99%)		146	3 (2.05%)	
17 (2015)	1026	865 (84.31%)		822	705 (85.77%)		204	160 (78.43%)	
18 (2017)	728	624 (85.71%)		502	440 (87.65%)		226	184 (81.42%)	
19 (2019)	453	165 (36.42%)		288	119 (41.32%)		165	46 (27.88%)	
Viral load (log10 copies/mL)			<1e-4			<1e-4			0.8
(3,4]	2101	895 (42.6%)		1748	723 (41.36%)		353	172 (48.73%)	
(4,5]	2231	1357 (60.82%)		1898	1180 (62.17%)		333	177 (53.15%)	
(5,Inf]	598	393 (65.72%)		508	348 (68.5%)		90	45 (50%)	
missing	1023	435 (42.52%)		1023	435 (42.52%)				

Percentages show the percentage of all samples within a category that have successful genotyping for all NNRTIs. *p*-values generated using the χ^2 test. PLWHIV=people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors

Supplementary Table 16: Number of viremic samples at participant visits that had successful genotyping for all NRTIs

	Viremic PLWHIV			Viremic pre-treatment PLWHIV			Viremic treatment-experienced PLWHIV		
	Participant-visits	**Genotype data for all NRTIs (%)**	***p*-value**	**Participant-visits**	**Genotype data for all NRTIs (%)**	***p*-value**	**Participant-visits**	**Genotype data for all NRTIs (%)**	***p*-value**
Overall	5953	3042 (51.1%)		5177	2659 (51.36%)		776	383 (49.36%)	
Age Category			0.38			0.48			0.81
(14,24]	1146	621 (54.19%)		1032	560 (54.26%)		114	61 (53.51%)	
(24,34]	2931	1490 (50.84%)		2573	1312 (50.99%)		358	178 (49.72%)	
(34,50]	1876	931 (49.63%)		1572	787 (50.06%)		304	144 (47.37%)	
Community type			0.93			0.97			0.29
Agrarian	1794	906 (50.5%)		1590	817 (51.38%)		204	89 (43.63%)	
Fishing	2861	1474 (51.52%)		2455	1255 (51.12%)		406	219 (53.94%)	
Trading	1298	662 (51%)		1132	587 (51.86%)		166	75 (45.18%)	
Sex			0.044			0.041			0.84
F	3188	1560 (48.93%)		2744	1344 (48.98%)		444	216 (48.65%)	
M	2765	1482 (53.6%)		2433	1315 (54.05%)		332	167 (50.3%)	
Survey round			<1e-4			<1e-4			<1e-4
15 (2012)	2281	1030 (45.16%)		2246	1029 (45.81%)		35	1 (2.86%)	
16 (2014)	1465	420 (28.67%)		1319	417 (31.61%)		146	3 (2.05%)	
17 (2015)	1026	817 (79.63%)		822	665 (80.9%)		204	152 (74.51%)	
18 (2017)	728	613 (84.2%)		502	431 (85.86%)		226	182 (80.53%)	
19 (2019)	453	162 (35.76%)		288	117 (40.62%)		165	45 (27.27%)	
Viral load (log10 copies/mL)			<1e-4			<1e-4			0.48
(3,4]	2101	844 (40.17%)		1748	684 (39.13%)		353	160 (45.33%)	
(4,5]	2231	1357 (60.82%)		1898	1180 (62.17%)		333	177 (53.15%)	
(5,Inf]	598	397 (66.39%)		508	351 (69.09%)		90	46 (51.11%)	
missing	1023	444 (43.4%)		1023	444 (43.4%)				

Percentages show the percentage of all samples within a category that have successful genotyping for all NRTIs. *p*-values generated using the χ^2 test. PLWHIV=people living with HIV. NRTI = nucleoside reverse transcriptase inhibitors

Supplementary Table 17: Number of viremic samples at participant visits that had successful genotyping for all PIs

	Viremic PLWHIV			Viremic pre-treatment PLWHIV			Viremic treatment-experienced PLWHIV		
	Participant-visits	Genotype data for all PIs (%)	*p*-value	Participant-visits	Genotype data for all PIs (%)	*p*-value	Participant-visits	Genotype data for all PIs (%)	*p*-value
Overall	5953	3567 (59.92%)		5177	3174 (61.31%)		776	393 (50.64%)	
Age Category			0.36			0.55			0.79
(14,24]	1146	725 (63.26%)		1032	662 (64.15%)		114	63 (55.26%)	
(24,34]	2931	1754 (59.84%)		2573	1572 (61.1%)		358	182 (50.84%)	
(34,50]	1876	1088 (58%)		1572	940 (59.8%)		304	148 (48.68%)	
Community type			0.39			0.62			0.18
Agrarian	1794	1037 (57.8%)		1590	946 (59.5%)		204	91 (44.61%)	
Fishing	2861	1765 (61.69%)		2455	1537 (62.61%)		406	228 (56.16%)	
Trading	1298	765 (58.94%)		1132	691 (61.04%)		166	74 (44.58%)	
Sex			0.16			0.18			0.93
F	3188	1857 (58.25%)		2744	1634 (59.55%)		444	223 (50.23%)	
M	2765	1710 (61.84%)		2433	1540 (63.3%)		332	170 (51.2%)	
Survey round			<1e-4			<1e-4			<1e-4
15 (2012)	2281	1408 (61.73%)		2246	1407 (62.64%)		35	1 (2.86%)	
16 (2014)	1465	502 (34.27%)		1319	499 (37.83%)		146	3 (2.05%)	
17 (2015)	1026	863 (84.11%)		822	706 (85.89%)		204	157 (76.96%)	
18 (2017)	728	625 (85.85%)		502	440 (87.65%)		226	185 (81.86%)	
19 (2019)	453	169 (37.31%)		288	122 (42.36%)		165	47 (28.48%)	
Viral load (log10 copies/mL)			<1e-4			<1e-4			0.7
(3,4]	2101	1079 (51.36%)		1748	910 (52.06%)		353	169 (47.88%)	
(4,5]	2231	1506 (67.5%)		1898	1328 (69.97%)		333	178 (53.45%)	
(5,Inf]	598	412 (68.9%)		508	366 (72.05%)		90	46 (51.11%)	
missing	1023	570 (55.72%)		1023	570 (55.72%)				

Percentages show the percentage of all samples within a category that have successful genotyping for all PIs. *p*-values generated using the χ^2 test. PLWHIV=people living with HIV. PI = protease inhibitor.

Supplementary Table 18: Number of viremic samples that had resistant genotypes for all drug classes

	Genotyped	Resistant (%)
INSTI	2593	19 (0.73%)
NNRTI	3080	381 (12.37%)
NRTI	3042	204 (6.71%)
PI	3567	66 (1.85%)

Supplementary Table 19: Number of viremic samples at participant visits that had successful genotyping for all NNRTIs, NRTIs, and PIs

	Viremic PLWHIV			Viremic pre-treatment PLWHIV			Viremic treatment-experienced PLWHIV		
	Participant-visits	**Genotype data for NNRTIs, NRTI, and PIs (%)**	***p*-value**	**Participant-visits**	**Genotype data for NNRTIs, NRTI, and PIs (%)**	***p*-value**	**Participant-visits**	**Genotype data for NNRTIs, NRTI, and PIs (%)**	***p*-value**
Overall	5953	2939 (49.37%)		5177	2565 (49.55%)		776	374 (48.2%)	
Age category			0.31			0.37			0.86
(14,24]	1146	603 (52.62%)		1032	544 (52.71%)		114	59 (51.75%)	
(24,34]	2931	1441 (49.16%)		2573	1268 (49.28%)		358	173 (48.32%)	
(34,50]	1876	895 (47.71%)		1572	753 (47.9%)		304	142 (46.71%)	
Community type			1			0.83			0.26
Agrarian	1794	884 (49.28%)		1590	796 (50.06%)		204	88 (43.14%)	
Fishing	2861	1413 (49.39%)		2455	1198 (48.8%)		406	215 (52.96%)	
Trading	1298	642 (49.46%)		1132	571 (50.44%)		166	71 (42.77%)	
Sex			0.042			0.039			0.85
F	3188	1506 (47.24%)		2744	1295 (47.19%)		444	211 (47.52%)	
M	2765	1433 (51.83%)		2433	1270 (52.2%)		332	163 (49.1%)	
Survey round			<1e-4			<1e-4			<1e-4
15 (2012)	2281	960 (42.09%)		2246	959 (42.7%)		35	1 (2.86%)	
16 (2014)	1465	403 (27.51%)		1319	400 (30.33%)		146	3 (2.05%)	
17 (2015)	1026	802 (78.17%)		822	658 (80.05%)		204	144 (70.59%)	
18 (2017)	728	612 (84.07%)		502	431 (85.86%)		226	181 (80.09%)	
19 (2019)	453	162 (35.76%)		288	117 (40.62%)		165	45 (27.27%)	
Viral load (log10 copies/mL)			<1e-4			<1e-4			0.33
(3,4]	2101	805 (38.32%)		1748	652 (37.3%)		353	153 (43.34%)	
(4,5]	2231	1324 (59.35%)		1898	1148 (60.48%)		333	176 (52.85%)	
(5,Inf]	598	390 (65.22%)		508	345 (67.91%)		90	45 (50%)	
missing	1023	420 (41.06%)		1023	420 (41.06%)				

Percentages show the percentage of all samples within a category that have successful genotyping for all INSTIs, NNRTI, NRTIs, and PIs. *p*-values generated using the χ^2 test. PLWHIV=people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = protease inhibitors.

Supplementary Table 20: Number of viremic samples at participant visits that had successful genotyping for all drugs

	Viremic PLWHIV			Viremic pre-treatment PLWHIV			Viremic treatment-experienced PLWHIV		
	Participant-visits	**Genotype data for all drug classes (%)**	***p*-value**	**Participant-visits**	**Genotype data for all drug classes (%)**	***p*-value**	**Participant-visits**	**Genotype data for all drug classes (%)**	***p*-value**
Overall	5953	2336 (39.24%)		5177	1965 (37.96%)		776	371 (47.81%)	
Age category			0.43			0.4			0.8
(14,24]	1146	479 (41.8%)		1032	420 (40.7%)		114	59 (51.75%)	
(24,34]	2931	1132 (38.62%)		2573	959 (37.27%)		358	173 (48.32%)	
(34,50]	1876	725 (38.65%)		1572	586 (37.28%)		304	139 (45.72%)	
Community type			0.84			0.29			0.19
Agrarian	1794	714 (39.8%)		1590	626 (39.37%)		204	88 (43.14%)	
Fishing	2861	1106 (38.66%)		2455	891 (36.29%)		406	215 (52.96%)	
Trading	1298	516 (39.75%)		1132	448 (39.58%)		166	68 (40.96%)	
Sex			0.014			0.0081			0.76
F	3188	1180 (37.01%)		2744	972 (35.42%)		444	208 (46.85%)	
M	2765	1156 (41.81%)		2433	993 (40.81%)		332	163 (49.1%)	
Survey round			<1e-4			<1e-4			<1e-4
15 (2012)	2281	494 (21.66%)		2246	494 (21.99%)		35	0 (0%)	
16 (2014)	1465	269 (18.36%)		1319	267 (20.24%)		146	2 (1.37%)	
17 (2015)	1026	799 (77.88%)		822	656 (79.81%)		204	143 (70.1%)	
18 (2017)	728	612 (84.07%)		502	431 (85.86%)		226	181 (80.09%)	
19 (2019)	453	162 (35.76%)		288	117 (40.62%)		165	45 (27.27%)	
Viral load (log10 copies/mL)			<1e-4			<1e-4			0.35
(3,4]	2101	656 (31.22%)		1748	504 (28.83%)		353	152 (43.06%)	
(4,5]	2231	1091 (48.9%)		1898	917 (48.31%)		333	174 (52.25%)	
(5,Inf]	598	356 (59.53%)		508	311 (61.22%)		90	45 (50%)	
missing	1023	233 (22.78%)		1023	233 (22.78%)				

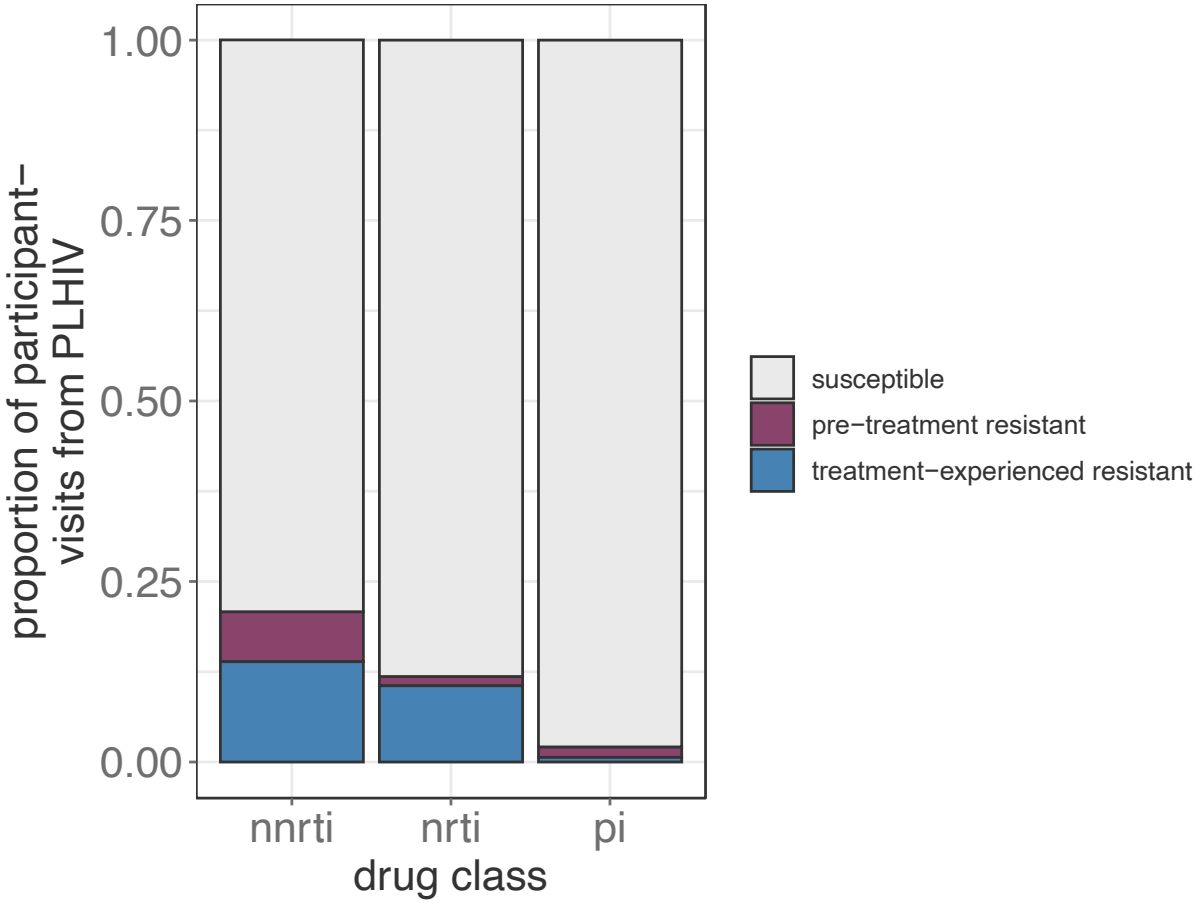
Percentages show the percentage of all samples within a category that have successful genotyping for all INSTIs, NNRTI, NRTIs, and PIs. *p*-values generated using the χ^2 test. PLWHIV=people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = protease inhibitors.

Supplementary Table 21: Population prevalence of viremic pre-treatment NNRTI, NRTI, and PI resistance

	NNRTI			NRTI			PI		
Survey round	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**
15 (2012)	0.65 (0.49, 0.87)	ref	ref	0.22 (0.14, 0.34)	ref	ref	0.27 (0.19, 0.4)	ref	ref
16 (2014)	0.54 (0.37, 0.79)	0.83 (0.51, 1.33)	0.44	0.11 (0.05, 0.26)	0.5 (0.19, 1.3)	0.15	0.14 (0.07, 0.29)	0.51 (0.22, 1.15)	0.11
17 (2015)	0.3 (0.23, 0.4)	0.47 (0.31, 0.69)	0.00017	0.08 (0.05, 0.13)	0.36 (0.18, 0.71)	0.0035	0.06 (0.03, 0.11)	0.2 (0.1, 0.43)	<0.0001
18 (2017)	0.25 (0.19, 0.33)	0.38 (0.25, 0.57)	<0.0001	0.04 (0.02, 0.09)	0.2 (0.09, 0.45)	0.0001	0.05 (0.03, 0.1)	0.19 (0.09, 0.39)	<0.0001
19 (2019)	0.13 (0.07, 0.24)	0.2 (0.1, 0.4)	<0.0001	0.02 (0.01, 0.1)	0.11 (0.03, 0.47)	0.0027	0.01 (0, 0.08)	0.04 (0.01, 0.31)	0.0018

Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. *p*-values that coefficients are different from 0 at the $\alpha = 0.05$ level were calculated using the Wald method. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = protease inhibitors.

Supplementary Figure 3: Prevalence of NNRTI, NRTI, and PI resistance among viremic PLHIV in survey round 18, stratified by treatment exposure



Supplementary Table 22: Population prevalence of viremic treatment-experienced NNRTI, NRTI, and PI resistance

	NNRTI			NRTI			PI		
Survey round	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**
17 (2015)	0.67 (0.54, 0.83)	ref	ref	0.59 (0.47, 0.75)	ref	ref	0.04 (0.02, 0.09)	ref	ref
18 (2017)	0.66 (0.53, 0.81)	0.98 (0.73, 1.31)	0.88	0.5 (0.4, 0.64)	0.85 (0.61, 1.18)	0.33	0.03 (0.01, 0.08)	0.8 (0.22, 3)	0.74

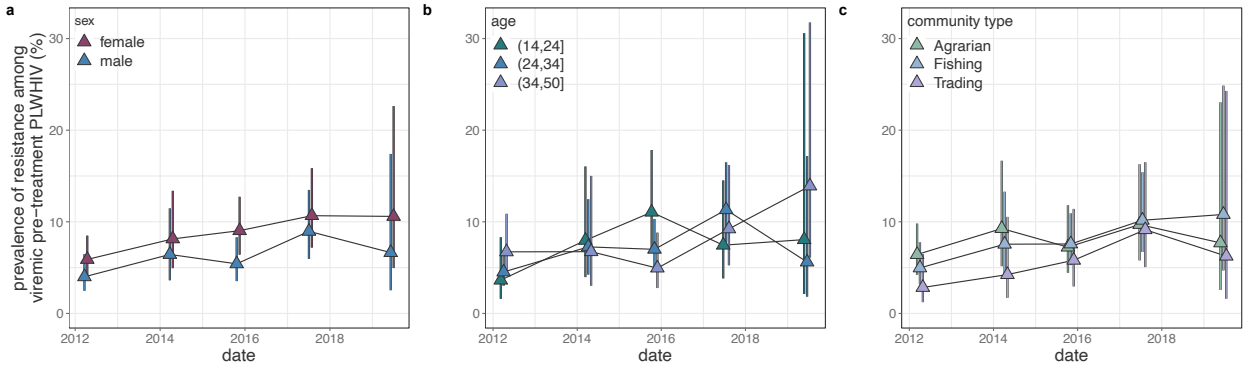
Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. *p*-values that coefficients are different from 0 at the $\alpha = 0.05$ level were calculated using the Wald method. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = integrase inhibitors.

Supplementary Table 23: Population prevalence of viremic NNRTI, NRTI, and PI resistance

Survey round	NNRTI			NRTI			PI		
	Prevalence (95% CI)	Prev. ratio (95% CI)	*p*-value	Prevalence (95% CI)	Prev. ratio (95% CI)	*p*-value	Prevalence (95% CI)	Prev. ratio (95% CI)	*p*-value
17 (2015)	0.96 (0.81, 1.13)	ref	ref	0.62 (0.5, 0.77)	ref	ref	0.1 (0.06, 0.17)	ref	ref
18 (2017)	0.9 (0.76, 1.06)	0.94 (0.74, 1.19)	0.6	0.51 (0.41, 0.64)	0.82 (0.6, 1.11)	0.2	0.09 (0.05, 0.15)	0.9 (0.42, 1.9)	0.77

Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. *p*-values that coefficients are different from 0 at the $\alpha = 0.05$ level were calculated using the Wald method. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = integrase inhibitors.

Supplementary Figure 4: Prevalence of NNRTI resistance among viremic pre-treatment PLHIV stratified by sex, age category, and community type



Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. For visual clarity, points are jittered along the x-axis. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors.

Supplementary Table 24: Prevalence of NNRTI resistance among viremic pre-treatment PLHIV stratified by sex, age category, and community type

	NNRTI Resistance among treatment-naïve PLHIV								
	R15 (2012)			R16 (2014)			R17 (2015)		
	Prevalence % (95% CI)	Prev. ratio (95% CI)	*p*-value	Prevalence % (95% CI)	Prev. ratio (95% CI)	*p*-value	Prevalence % (95% CI)	Prev. ratio (95% CI)	*p*-value
Age category									
(14,24]	3.64 (1.59, 8.32)	ref	ref	7.98 (3.97, 16.04)	2.19 (0.74, 6.47)	0.16	11.05 (6.84, 17.84)	3.04 (1.17, 7.89)	0.023
(24,34]	4.52 (3.02, 6.77)	1.24 (0.5, 3.12)	0.64	7.27 (4.24, 12.46)	0.73 (0.2, 2.62)	0.63	7 (4.75, 10.31)	0.51 (0.17, 1.54)	0.23
(34,50]	6.72 (4.16, 10.88)	1.85 (0.71, 4.81)	0.21	6.74 (3.02, 15.01)	0.46 (0.11, 1.91)	0.28	4.95 (2.78, 8.82)	0.24 (0.07, 0.82)	0.022
Community type									
Agrarian	6.43 (4.21, 9.81)	ref	ref	9.27 (5.15, 16.66)	1.44 (0.7, 2.97)	0.32	7.24 (4.44, 11.81)	1.13 (0.59, 2.15)	0.72
Fishing	4.97 (3.18, 7.76)	0.77 (0.42, 1.43)	0.41	7.56 (4.3, 13.29)	1.05 (0.38, 2.92)	0.92	7.59 (5.27, 10.92)	1.36 (0.57, 3.22)	0.49
Trading	2.84 (1.24, 6.48)	0.44 (0.17, 1.12)	0.084	4.25 (1.71, 10.54)	1.04 (0.25, 4.32)	0.96	5.81 (2.96, 11.4)	1.82 (0.52, 6.33)	0.35
Sex									
F	5.9 (4.1, 8.49)	ref	ref	8.13 (4.94, 13.39)	1.38 (0.74, 2.56)	0.31	9.04 (6.41, 12.73)	1.53 (0.93, 2.53)	0.095
M	4 (2.47, 6.47)	0.68 (0.37, 1.24)	0.21	6.44 (3.62, 11.47)	1.17 (0.44, 3.09)	0.76	5.42 (3.53, 8.3)	0.88 (0.39, 2)	0.77

	NNRTI Resistance among treatment-naïve PLHIV					
	R18 (2017)			R19 (2019)		
	Prevalence % (95% CI)	Prev. ratio (95% CI)	*p*-value	Prevalence % (95% CI)	Prev. ratio (95% CI)	*p*-value
Age category						
(14,24]	7.45 (3.83, 14.52)	2.05 (0.71, 5.93)	0.19	8.06 (2.12, 30.58)	2.21 (0.46, 10.63)	0.32
(24,34]	11.31 (7.76, 16.49)	1.22 (0.37, 4.04)	0.74	5.61 (1.84, 17.18)	0.56 (0.08, 4.02)	0.56
(34,50]	9.23 (5.26, 16.2)	0.67 (0.18, 2.45)	0.54	13.91 (6.09, 31.75)	0.93 (0.15, 5.87)	0.94
Community type						
Agrarian	9.7 (5.79, 16.28)	1.51 (0.77, 2.94)	0.23	7.7 (2.57, 23.05)	1.2 (0.37, 3.88)	0.76
Fishing	10.17 (6.71, 15.41)	1.35 (0.55, 3.35)	0.51	10.8 (4.69, 24.86)	1.81 (0.4, 8.19)	0.44
Trading	9.14 (5.06, 16.5)	2.13 (0.63, 7.2)	0.22	6.25 (1.61, 24.28)	1.84 (0.26, 13.27)	0.55
Sex						
F	10.67 (7.19, 15.85)	1.81 (1.06, 3.1)	0.031	10.59 (4.95, 22.62)	1.79 (0.77, 4.17)	0.17

M	8.96 (5.96, 13.48)	1.24 (0.54, 2.83)	0.61	6.64 (2.53, 17.41)	0.92 (0.24, 3.63)	0.91

Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. *p*-values that coefficients are different from 0 at the $\alpha = 0.05$ level were calculated using the Wald method. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors.

Supplementary Table 25: Prevalence of NNRTI resistance among seroincident, pre-treatment PLHIV

	NNRTI		
Survey round	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**
15 (2012)	9.97 (3, 33.13)	ref	ref
16 (2014)	11.03 (4.27, 28.46)	1.11 (0.24, 5.11)	0.9
17 (2015)	9.9 (4.55, 21.52)	0.99 (0.24, 4.15)	0.99
18 (2017)	9.95 (4.65, 21.28)	1 (0.24, 4.14)	1
19 (2019)	8.95 (2.39, 33.51)	0.9 (0.15, 5.35)	0.91

Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. *p*-values that coefficients are different from 0 at the $\alpha = 0.05$ level were calculated using the Wald method. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors.

Seroincident sub-analysis power calculation

Based on the population-level analysis, we estimate that 9.77% of pre-treatment PLHIV in round 18 had NNRTI-resistant genotypes compared to 4.98% in round 15.

Assuming normality, we let $p_1 = 0.0498$ and $p_2 = 0.0977$. At the $\alpha = 0.05$, $\beta = 0.80$ level we can calculate the number of samples we would need at each time point to have sufficient power to detect this difference using the equation[3]:

$$n = (Z_{\frac{\alpha}{2}} + Z_{\frac{\beta}{2}})^2 \frac{(p_1(1-p_1) + p_2(1-p_2))}{(p_1 - p_2)^2}$$

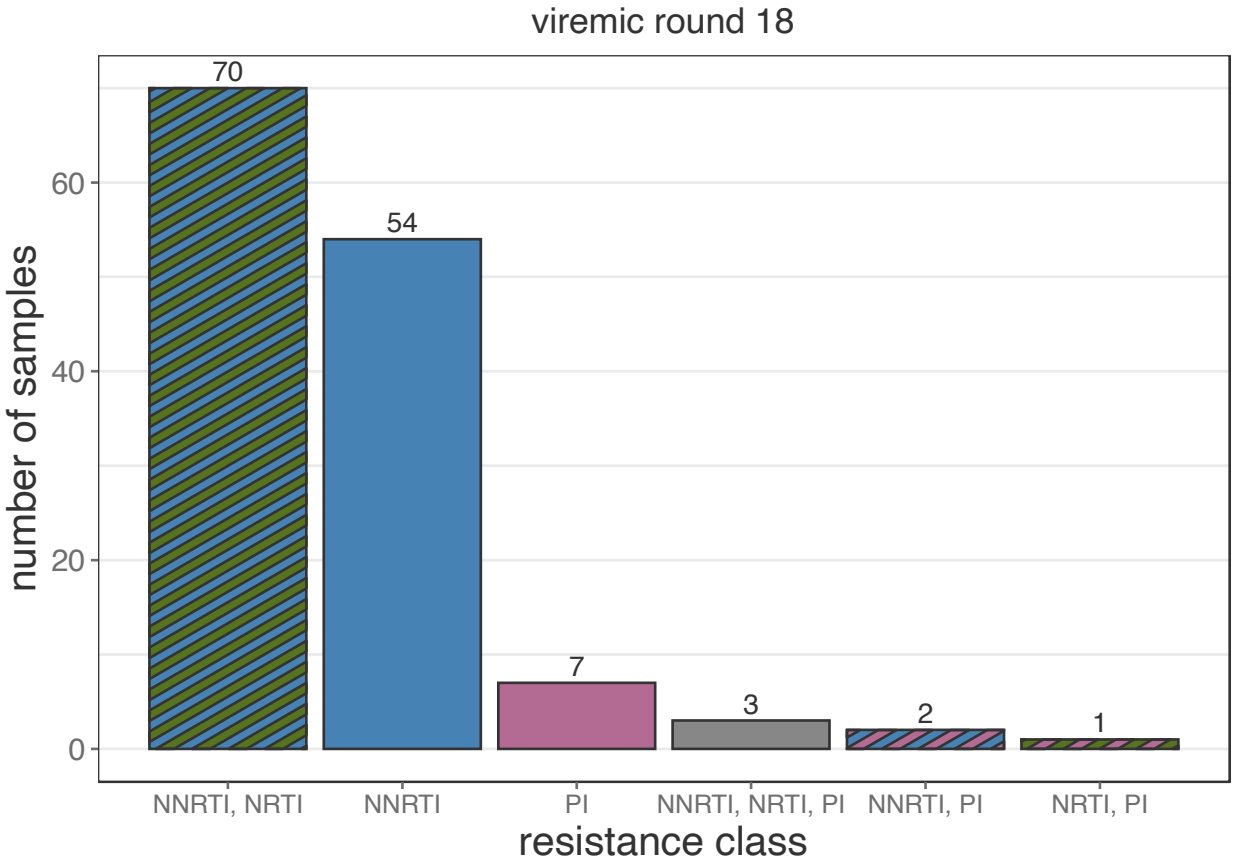
We thus calculate that we would need a sample size of $n = 432$ at each time point. In the seroincident analysis in Supplementary Table 17 we have data from 39, 32, 67, 59, and 26 PLHIV in rounds 15 through 19, respectively.

Supplementary Table 26: Prevalence of NNRTI, NRTI, and PI resistance among viremic treatment-experienced PLHIV

	NNRTI			NRTI			PI		
Survey round	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**
17 (2015)	52.47 (45.17, 60.94)	ref	ref	46.44 (38.99, 55.3)	ref	ref	2.92 (1.21, 7.07)	ref	ref
18 (2017)	47.67 (40.94, 55.5)	0.91 (0.73, 1.12)	0.38	36.55 (30.14, 44.31)	0.79 (0.61, 1.02)	0.071	2.18 (0.83, 5.76)	0.75 (0.2, 2.77)	0.66

Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. *p*-values that coefficients are different from 0 at the $\alpha = 0.05$ level were calculated using the Wald method. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = integrase inhibitors.

Supplementary Figure 5: Patterns of multi-class resistance in Rakai Community Cohort Study, 2017



Multi-class resistance patterns among 137 round 18 participant visits with resistance prediction for all NNRTIs, NRTIs, and PIs with resistance to at least one of the three classes. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = integrase inhibitors.

Supplementary Table 27: Round 18 prevalence of NNRTI single- and multi-class resistance among PLHIV

	Prevalence % (95% CI)
NNRTI single-class resistance	8.9 (6.9, 11.49)
NNRTI multi-class resistance	12.04 (9.73, 14.89)

Supplementary Table 28: Population prevalence of NNRTI, NRTI, and PI resistance in survey round 18 stratified by community type

		NNRTI		NRTI		PI	
Community type	Viremia (%)	Observed % (95% CI)	Expected % (95% CI)	Observed % (95% CI)	Expected % (95% CI)	Observed % (95% CI)	Expected % (95% CI)
Agrarian	2.4 (2.1, 2.74)	0.49 (0.35, 0.69)	0.5 (0.43, 0.58)	0.27 (0.18, 0.43)	0.28 (0.23, 0.35)	0.01 (0, 0.1)	0.05 (0.03, 0.09)
Fishing	7.42 (6.71, 8.2)	2.35 (1.86, 2.97)	1.54 (1.32, 1.8)	1.46 (1.08, 1.96)	0.88 (0.71, 1.09)	0.28 (0.14, 0.55)	0.16 (0.09, 0.27)
Trading	2.62 (2.25, 3.04)	0.59 (0.41, 0.85)	0.54 (0.47, 0.64)	0.27 (0.16, 0.46)	0.31 (0.25, 0.38)	0.08 (0.03, 0.22)	0.05 (0.03, 0.09)

Estimates generated using robust log-Poisson regression. 95% confidence intervals around the mean value generated using the Wald method. Expected values generated using the mean and 95% CI for the prevalence of resistance among viremic PLHIV across all community types extrapolated to the estimated mean prevalence of viremia within individual community types. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = protease inhibitors.

Supplementary Table 29: Population prevalence of NNRTI, NRTI, and PI resistance in survey round 18 stratified by community

Community			NNRTI		NRTI		PI	
Type	Number	Viremia (%)	Observed % (95% CI)	Expected % (95% CI)	Observed % (95% CI)	Expected % (95% CI)	Observed % (95% CI)	Expected % (95% CI)
Fishing	774	8.19 (5.91, 11.36)	2.69 (1.23, 5.88)	1.7 (1.46, 1.99)	1.35 (0.44, 4.14)	0.97 (0.78, 1.2)	0 (0, 0)	0.17 (0.1, 0.29)
Fishing	770	7.54 (5.69, 9.98)	1.9 (0.92, 3.94)	1.57 (1.35, 1.83)	1.41 (0.6, 3.35)	0.89 (0.72, 1.11)	0.28 (0.04, 1.94)	0.16 (0.09, 0.27)
Fishing	38	7.46 (6.54, 8.51)	2.19 (1.6, 3.01)	1.55 (1.33, 1.81)	1.3 (0.86, 1.97)	0.88 (0.71, 1.1)	0.24 (0.09, 0.63)	0.16 (0.09, 0.27)
Fishing	771	6.92 (5.52, 8.67)	2.97 (1.89, 4.66)	1.44 (1.24, 1.68)	1.97 (1.13, 3.43)	0.82 (0.66, 1.02)	0.5 (0.16, 1.53)	0.14 (0.08, 0.25)
Agrarian	58	5.22 (2.39, 11.37)	0 (0, 0)	1.09 (0.93, 1.27)	0 (0, 0)	0.62 (0.5, 0.77)	0 (0, 0)	0.11 (0.06, 0.19)
Agrarian	74	4.83 (2.34, 9.94)	0.94 (0.14, 6.54)	1 (0.86, 1.17)	0.99 (0.14, 6.84)	0.57 (0.46, 0.71)	0 (0, 0)	0.1 (0.06, 0.17)
Trading	772	4.81 (2.54, 9.1)	0 (0, 0)	1 (0.86, 1.17)	0 (0, 0)	0.57 (0.46, 0.71)	0 (0, 0)	0.1 (0.06, 0.17)
Trading	391	4.63 (2.72, 7.87)	0.46 (0.07, 3.22)	0.96 (0.83, 1.12)	0 (0, 0)	0.55 (0.44, 0.68)	0 (0, 0)	0.1 (0.06, 0.17)
Agrarian	370	4.12 (2.37, 7.18)	0.46 (0.07, 3.19)	0.86 (0.74, 1)	0 (0, 0)	0.49 (0.39, 0.61)	0 (0, 0)	0.09 (0.05, 0.15)
Agrarian	773	4.1 (2.57, 6.52)	0.32 (0.05, 2.2)	0.85 (0.73, 0.99)	0.33 (0.05, 2.27)	0.49 (0.39, 0.6)	0 (0, 0)	0.09 (0.05, 0.15)
Trading	8	4.09 (2.49, 6.71)	1.4 (0.53, 3.67)	0.85 (0.73, 0.99)	1.03 (0.34, 3.14)	0.48 (0.39, 0.6)	0.34 (0.05, 2.38)	0.09 (0.05, 0.15)
Trading	1	4 (1.53, 10.45)	2.75 (0.71, 10.67)	0.83 (0.71, 0.97)	1.52 (0.22, 10.44)	0.47 (0.38, 0.59)	1.43 (0.21, 9.86)	0.08 (0.05, 0.14)
Agrarian	77	3.91 (2.13, 7.17)	3.01 (1.38, 6.58)	0.81 (0.7, 0.95)	0.99 (0.25, 3.88)	0.46 (0.37, 0.57)	0 (0, 0)	0.08 (0.05, 0.14)
Agrarian	23	3.47 (2.21, 5.47)	0.74 (0.24, 2.27)	0.72 (0.62, 0.84)	0.74 (0.24, 2.26)	0.41 (0.33, 0.51)	0 (0, 0)	0.07 (0.04, 0.12)
Trading	4	3.16 (1.85, 5.4)	0.65 (0.17, 2.57)	0.66 (0.56, 0.77)	0.32 (0.05, 2.22)	0.37 (0.3, 0.46)	0 (0, 0)	0.07 (0.04, 0.11)
Agrarian	106	3 (1.9, 4.72)	0.21 (0.03, 1.5)	0.62 (0.53, 0.73)	0.22 (0.03, 1.51)	0.35 (0.29, 0.44)	0 (0, 0)	0.06 (0.04, 0.11)
Trading	34	2.73 (1.81, 4.12)	0.66 (0.25, 1.73)	0.57 (0.49, 0.66)	0.16 (0.02, 1.12)	0.32 (0.26, 0.4)	0 (0, 0)	0.06 (0.03, 0.1)
Agrarian	602	2.69 (1.84, 3.92)	0.91 (0.44, 1.89)	0.56 (0.48, 0.65)	0.64 (0.27, 1.53)	0.32 (0.26, 0.39)	0.13 (0.02, 0.92)	0.06 (0.03, 0.1)
Agrarian	29	2.6 (1.31, 5.15)	0.84 (0.21, 3.32)	0.54 (0.46, 0.63)	0.41 (0.06, 2.84)	0.31 (0.25, 0.38)	0 (0, 0)	0.05 (0.03, 0.09)
Agrarian	108	2.56 (1.5, 4.39)	0.51 (0.13, 2.01)	0.53 (0.46, 0.62)	0.24 (0.03, 1.7)	0.3 (0.24, 0.38)	0 (0, 0)	0.05 (0.03, 0.09)
Trading	22	2.47 (0.94, 6.5)	0.8 (0.12, 5.58)	0.51 (0.44, 0.6)	0.78 (0.11, 5.41)	0.29 (0.24, 0.36)	0 (0, 0)	0.05 (0.03, 0.09)
Trading	24	2.38 (1.6, 3.53)	0.52 (0.2, 1.38)	0.5 (0.42, 0.58)	0.38 (0.12, 1.16)	0.28 (0.23, 0.35)	0 (0, 0)	0.05 (0.03, 0.09)
Trading	754	2.38 (1.39, 4.07)	0.48 (0.12, 1.88)	0.49 (0.42, 0.58)	0 (0, 0)	0.28 (0.23, 0.35)	0.23 (0.03, 1.64)	0.05 (0.03, 0.09)
Agrarian	16	2.33 (1.06, 5.15)	0.48 (0.07, 3.37)	0.49 (0.42, 0.57)	0 (0, 0)	0.28 (0.22, 0.34)	0 (0, 0)	0.05 (0.03, 0.08)

Trading	7	2.32 (1.36, 3.97)	0.46 (0.12, 1.83)	0.48 (0.41, 0.56)	0.24 (0.03, 1.69)	0.27 (0.22, 0.34)	0 (0, 0)	0.05 (0.03, 0.08)
Trading	2	2.24 (1.28, 3.92)	1.01 (0.38, 2.65)	0.47 (0.4, 0.54)	0.24 (0.03, 1.68)	0.27 (0.21, 0.33)	0 (0, 0)	0.05 (0.03, 0.08)
Trading	33	2.19 (0.92, 5.22)	0 (0, 0)	0.46 (0.39, 0.53)	0 (0, 0)	0.26 (0.21, 0.32)	0 (0, 0)	0.05 (0.03, 0.08)
Agraria n	19	2.13 (1.33, 3.41)	0.32 (0.08, 1.26)	0.44 (0.38, 0.52)	0.16 (0.02, 1.09)	0.25 (0.2, 0.31)	0 (0, 0)	0.04 (0.03, 0.08)
Agraria n	6	1.95 (1.02, 3.72)	0.29 (0.04, 2.04)	0.41 (0.35, 0.47)	0 (0, 0)	0.23 (0.19, 0.29)	0 (0, 0)	0.04 (0.02, 0.07)
Trading	89	1.89 (0.85, 4.17)	0 (0, 0)	0.39 (0.34, 0.46)	0 (0, 0)	0.22 (0.18, 0.28)	0.42 (0.06, 2.9)	0.04 (0.02, 0.07)
Agraria n	57	1.85 (0.78, 4.41)	0.45 (0.06, 3.12)	0.39 (0.33, 0.45)	0 (0, 0)	0.22 (0.18, 0.27)	0 (0, 0)	0.04 (0.02, 0.07)
Agraria n	56	1.8 (0.68, 4.76)	0.57 (0.08, 3.97)	0.38 (0.32, 0.44)	0.54 (0.08, 3.8)	0.21 (0.17, 0.26)	0 (0, 0)	0.04 (0.02, 0.06)
Trading	40	1.59 (0.89, 2.86)	0.19 (0.03, 1.34)	0.33 (0.28, 0.39)	0 (0, 0)	0.19 (0.15, 0.23)	0 (0, 0)	0.03 (0.02, 0.06)
Agraria n	55	1.59 (0.8, 3.16)	0.25 (0.04, 1.71)	0.33 (0.28, 0.39)	0 (0, 0)	0.19 (0.15, 0.23)	0 (0, 0)	0.03 (0.02, 0.06)
Agraria n	36	1.59 (0.76, 3.32)	0 (0, 0)	0.33 (0.28, 0.39)	0 (0, 0)	0.19 (0.15, 0.23)	0 (0, 0)	0.03 (0.02, 0.06)
Agraria n	62	1.56 (0.59, 4.13)	1.09 (0.28, 4.27)	0.33 (0.28, 0.38)	0.58 (0.08, 4.07)	0.19 (0.15, 0.23)	0 (0, 0)	0.03 (0.02, 0.06)
Agraria n	5	1.44 (0.72, 2.86)	0 (0, 0)	0.3 (0.26, 0.35)	0 (0, 0)	0.17 (0.14, 0.21)	0 (0, 0)	0.03 (0.02, 0.05)
Trading	94	1.01 (0.14, 7.1)	1.4 (0.2, 9.68)	0.21 (0.18, 0.25)	1.51 (0.22, 10.35)	0.12 (0.1, 0.15)	0 (0, 0)	0.02 (0.01, 0.04)
Agraria n	107	0.97 (0.37, 2.57)	0.29 (0.04, 2.04)	0.2 (0.17, 0.23)	0.28 (0.04, 1.96)	0.11 (0.09, 0.14)	0 (0, 0)	0.02 (0.01, 0.03)
Agraria n	401	0.87 (0.22, 3.46)	0 (0, 0)	0.18 (0.16, 0.21)	0 (0, 0)	0.1 (0.08, 0.13)	0 (0, 0)	0.02 (0.01, 0.03)
Agraria n	120	0.46 (0.06, 3.24)	0 (0, 0)	0.1 (0.08, 0.11)	0 (0, 0)	0.05 (0.04, 0.07)	0 (0, 0)	0.01 (0.01, 0.02)

Estimates generated using robust log-Poisson regression. 95% confidence intervals around the mean value generated using the Wald method. Expected values generated using the mean and 95% CI for the prevalence of resistance among viremic PLHIV across all community types extrapolated to the estimated mean prevalence of viremia within individual community types. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors. PI = protease inhibitors.

Supplementary Table 30: Prevalence of amino acid substitutions among viremic pre-treatment PLHIV

Substitution	R15 (2012)			R16 (2014)			R17 (2015)		
	Prevalence % (95% CI)	Prev. ratio (95% CI)	*p*-value	Prevalence % (95% CI)	Prev. ratio (95% CI)	*p*-value	Prevalence % (95% CI)	Prev. ratio (95% CI)	*p*-value
inT97A	7.83 (5.66, 10.83)	ref	<0.001	10.18 (6.61, 15.69)	1.3 (0.76, 2.23)	0.34	9.32 (7.33, 11.85)	1.19 (0.8, 1.78)	0.4
rtK103N	3.08 (1.61, 5.89)	ref	<0.001	4.54 (2.35, 8.79)	1.48 (0.58, 3.73)	0.41	3.97 (2.72, 5.81)	1.29 (0.61, 2.74)	0.51
rtE138A	1.47 (0.75, 2.86)	ref	<0.001	2.04 (0.93, 4.45)	1.38 (0.5, 3.87)	0.53	3.63 (2.46, 5.36)	2.47 (1.14, 5.35)	0.022
rtM41L	0.3 (0.11, 0.82)	ref	<0.001	0.54 (0.13, 2.16)	1.8 (0.33, 10.01)	0.5	2.67 (1.69, 4.23)	8.95 (2.96, 27.03)	<0.001
rtV179D	0.24 (0.05, 1.05)	ref	<0.001	0.42 (0.09, 1.89)	1.78 (0.21, 14.71)	0.59	1.21 (0.58, 2.54)	5.1 (0.97, 26.86)	0.055
prL33F	0.63 (0.22, 1.76)	ref	<0.001	1.35 (0.46, 4)	2.16 (0.48, 9.66)	0.31	1.05 (0.47, 2.34)	1.67 (0.45, 6.18)	0.44
rtK101E	0.66 (0.25, 1.77)	ref	<0.001	0.86 (0.24, 3.12)	1.3 (0.26, 6.53)	0.75	1 (0.48, 2.09)	1.51 (0.44, 5.14)	0.51
rtE138G	0.58 (0.19, 1.85)	ref	<0.001	1.47 (0.37, 5.82)	2.51 (0.42, 15.1)	0.31	0.27 (0.07, 1.08)	0.46 (0.08, 2.8)	0.4
prQ58E	0.34 (0.09, 1.38)	ref	<0.001	0.51 (0.07, 3.61)	1.48 (0.13, 16.33)	0.75	0.27 (0.07, 1.09)	0.79 (0.11, 5.63)	0.81
rtV108I	0.75 (0.26, 2.14)	ref	<0.001	0.93 (0.27, 3.18)	1.24 (0.25, 6.28)	0.79	0.86 (0.36, 2.07)	1.15 (0.29, 4.54)	0.84
rtG190A	1.04 (0.42, 2.58)	ref	<0.001	0.29 (0.04, 2.06)	0.28 (0.03, 2.43)	0.25	0.7 (0.29, 1.69)	0.68 (0.19, 2.4)	0.54
rtM184V	0.28 (0.07, 1.18)	ref	<0.001	0.43 (0.11, 1.71)	1.5 (0.21, 10.95)	0.69	1.3 (0.68, 2.51)	4.59 (0.96, 21.99)	0.056
rtV106I	0.92 (0.38, 2.23)	ref	<0.001	1.42 (0.47, 4.26)	1.54 (0.38, 6.35)	0.55	1.18 (0.59, 2.36)	1.29 (0.42, 3.96)	0.66
inE157Q	0.26 (0.06, 1.06)	ref	<0.001	0 (0, 0)	0 (0, 0)	<0.001	0.56 (0.21, 1.49)	2.16 (0.39, 12.03)	0.38
prV32I	1.42 (0.6, 3.37)	ref	<0.001	0.08 (0.01, 0.6)	0.06 (0.01, 0.51)	0.0099	0.76 (0.31, 1.84)	0.54 (0.15, 1.85)	0.32

Substitution	R18 (2017)			R19 (2019)		
	Prevalence % (95% CI)	Prev. ratio (95% CI)	*p*-value	Prevalence % (95% CI)	Prev. ratio (95% CI)	*p*-value
inT97A	10.33 (7.83, 13.63)	1.32 (0.86, 2.02)	0.2	20.95 (14.63, 30.01)	2.68 (1.65, 4.34)	<0.001
rtK103N	7.27 (5.17, 10.2)	2.36 (1.13, 4.92)	0.022	5.95 (2.87, 12.34)	1.93 (0.73, 5.14)	0.19
rtE138A	2.8 (1.6, 4.89)	1.91 (0.8, 4.55)	0.15	2.67 (0.87, 8.15)	1.82 (0.49, 6.67)	0.37
rtM41L	2.23 (1.2, 4.12)	7.47 (2.3, 24.26)	<0.001	6.33 (3.06, 13.09)	21.21 (6.14, 73.29)	<0.001
rtV179D	1.62 (0.77, 3.38)	6.82 (1.3, 35.86)	0.023	1.45 (0.35, 6.02)	6.11 (0.78, 47.89)	0.085
prL33F	1.57 (0.75, 3.28)	2.51 (0.7, 8.91)	0.16	1.86 (0.47, 7.34)	2.97 (0.53, 16.54)	0.21

rtK101E	1.12 (0.47, 2.69)	1.69 (0.45, 6.28)	0.43	2.85 (0.93, 8.74)	4.29 (0.97, 19.01)	0.056
rtE138G	0.93 (0.35, 2.46)	1.58 (0.35, 7.17)	0.55	0 (0, 0)	0 (0, 0)	<0.001
prQ58E	0.92 (0.35, 2.45)	2.69 (0.49, 14.66)	0.25	1.87 (0.47, 7.41)	5.42 (0.77, 38.38)	0.09
rtV108I	0.91 (0.34, 2.41)	1.21 (0.29, 5.11)	0.79	1.87 (0.48, 7.39)	2.51 (0.45, 14.15)	0.3
rtG190A	0.9 (0.34, 2.4)	0.87 (0.23, 3.31)	0.84	0 (0, 0)	0 (0, 0)	<0.001
rtM184V	0.9 (0.34, 2.39)	3.17 (0.56, 17.8)	0.19	0.95 (0.13, 6.66)	3.34 (0.3, 37.32)	0.33
rtV106I	0.71 (0.23, 2.21)	0.78 (0.19, 3.27)	0.73	0 (0, 0)	0 (0, 0)	<0.001
inE157Q	0.7 (0.23, 2.16)	2.69 (0.44, 16.37)	0.28	0.93 (0.13, 6.52)	3.58 (0.32, 39.67)	0.3
prV32I	0.69 (0.22, 2.12)	0.48 (0.12, 2.01)	0.32	0.92 (0.13, 6.45)	0.65 (0.08, 5.47)	0.69

Only the 15 most prevalent substitutions in round 18 are shown. Estimates generated using robust log-Poisson regression. 95% confidence intervals around the mean value generated using the Wald method. Expected values generated using the mean and 95% CI for the prevalence of resistance among viremic PLHIV across all community types extrapolated to the estimated mean prevalence of viremia within individual community types. PLHIV = people living with HIV. RT = reverse transcriptase. IN = integrase. PR = protease.

Supplementary Table 31: Prevalence of rtK103N among viremic pre-treatment PLHIV stratified by sex

	R15 (2012)			R16 (2014)			R17 (2015)		
	Prevalence % (95% CI)	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**
Sex									
F	3.92 (1.65, 9.31)	ref	ref	4.64 (1.8, 12)	1.19 (0.33, 4.29)	0.8	4.92 (2.99, 8.09)	1.26 (0.46, 3.41)	0.65
M	2.18 (0.89, 5.35)	0.56 (0.16, 1.94)	0.36	4.45 (1.78, 11.12)	1.72 (0.28, 10.61)	0.56	3.14 (1.74, 5.65)	1.15 (0.26, 4.98)	0.85

	R18 (2017)			R19 (2019)		
	Prevalence % (95% CI)	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**
Sex						
F	9.04 (5.82, 14.07)	2.31 (0.87, 6.1)	0.091	6.91 (2.67, 17.88)	1.76 (0.49, 6.38)	0.39
M	5.71 (3.37, 9.69)	1.14 (0.27, 4.73)	0.86	4.82 (1.58, 14.75)	1.26 (0.18, 8.62)	0.82

Supplementary Table 32: Prevalence of amino acid substitutions among viremic treatment-experienced PLHIV

	R17 (2015)	R18 (2017)
Substitution	**Prevalence % (95% CI)**	**Prevalence % (95% CI)**
rtM184V	38.08 (30.77, 47.13)	33.95 (27.67, 41.65)
rtK103N	27.91 (21.41, 36.38)	26.73 (20.93, 34.14)
rtY181C	20.17 (14.57, 27.92)	14.08 (9.84, 20.15)
rtK65R	14.46 (9.56, 21.88)	12.61 (8.52, 18.65)
rtG190A	12.77 (8.31, 19.65)	12 (8.09, 17.79)
inT97A	10.59 (6.31, 17.78)	9.98 (6.43, 15.51)
rtK101E	10.49 (6.44, 17.07)	9.49 (6.03, 14.94)
rtK219E	7.09 (3.86, 13.05)	8.04 (4.85, 13.31)
rtV108I	4.42 (2.21, 8.87)	7.36 (4.35, 12.45)
rtH221Y	7.97 (4.49, 14.18)	6.99 (4.13, 11.85)
rtN348I	9.28 (5.57, 15.45)	6.53 (3.77, 11.32)
rtK70R	5.87 (3.05, 11.31)	6.3 (3.55, 11.19)
rtM41L	6.3 (3.38, 11.73)	5.55 (3.03, 10.15)
rtP225H	3.79 (1.66, 8.67)	5.5 (3, 10.09)
rtD67N	4.15 (1.85, 9.34)	5.45 (2.97, 10)

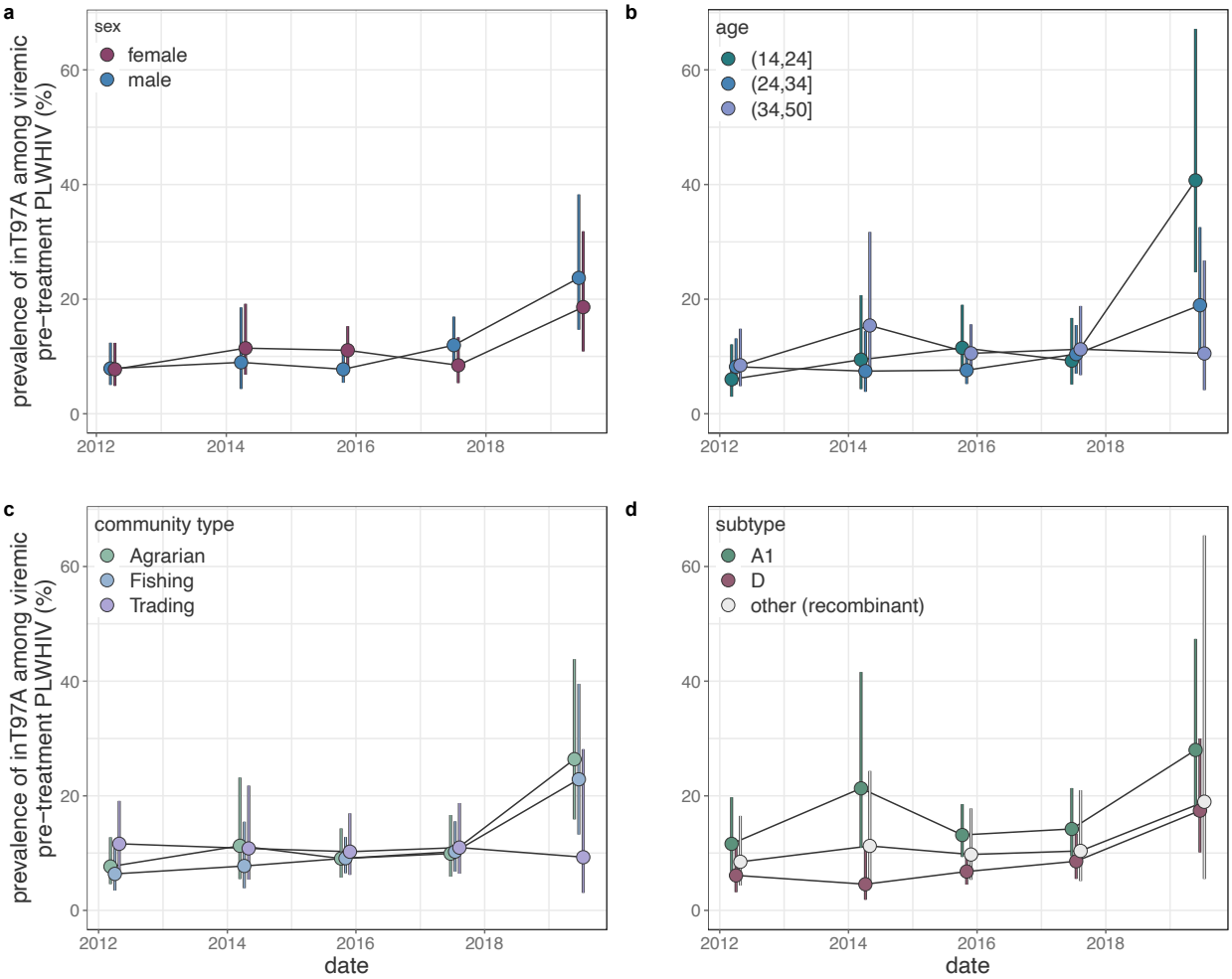
Only the 15 most prevalent substitutions in round 18 are shown. Estimates generated using robust log-Poisson regression. 95% confidence intervals around the mean value generated using the Wald method. Expected values generated using the mean and 95% CI for the prevalence of resistance among viremic PLHIV across all community types extrapolated to the estimated mean prevalence of viremia within individual community types. PLHIV = people living with HIV. RT = reverse transcriptase. IN = integrase. PR = protease.

Supplementary Table 33: Prevalence of inT97A among viremic pre-treatment PLHIV by correlation structure

	Independence			Exchangeable			AR1			Unstructured		
Survey round	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**
15 (2012)	7.83 (5.66, 10.83)	ref	ref	7.91 (5.76, 10.86)	ref	ref	7.9 (5.75, 10.85)	ref	ref	7.91 (5.76, 10.86)	ref	ref
16 (2014)	10.18 (6.61, 15.69)	1.3 (0.76, 2.23)	0.34	10.62 (7.06, 15.97)	1.34 (0.8, 2.24)	0.26	10.67 (7.09, 16.06)	1.35 (0.81, 2.26)	0.25	10.62 (7.08, 15.94)	1.34 (0.81, 2.24)	2.60E -01
17 (2015)	9.32 (7.33, 11.85)	1.19 (0.8, 1.76)	0.38	9.54 (7.62, 11.95)	1.21 (0.83, 1.76)	0.33	9.49 (7.57, 11.9)	1.2 (0.82, 1.76)	0.34	9.65 (7.74, 12.04)	1.22 (0.84, 1.78)	0.3
18 (2017)	10.33 (7.83, 13.63)	1.32 (0.86, 2.02)	0.2	10.42 (7.99, 13.59)	1.32 (0.87, 1.99)	0.19	10.34 (7.91, 13.52)	1.31 (0.87, 1.98)	0.2	10.14 (7.71, 13.34)	1.28 (0.85, 1.95)	0.24
19 (2019)	20.95 (14.63, 30.01)	2.68 (1.65, 4.34)	<0.00 01	20.6 (14.52, 29.21)	2.6 (1.63, 4.17)	<0.00 01	20.48 (14.4, 29.13)	2.59 (1.61, 4.17)	<0.00 01	20.39 (14.3, 29.07)	2.58 (1.6, 4.15)	<0.00 01

Estimates were generated using robust log-Poisson regression with generalized estimating equation. PLHIV = people living with HIV. NNRTI = non-nucleoside reverse transcriptase inhibitors. NRTI = nucleoside reverse transcriptase inhibitors.

Supplementary Figure 6: Prevalence of inT97A among viremic pre-treatment PLHIV stratified by sex, age category, community type, and subtype



Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. Only the three most common subtypes are shown. Subtypes were inferred by identifying the closest reference match to each sequenced sample. For visual clarity, points are jittered along the x-axis.

Supplementary Table 34: Prevalence of inT97A among viremic pre-treatment PLHIV stratified by sex, age category, community type, and subtype

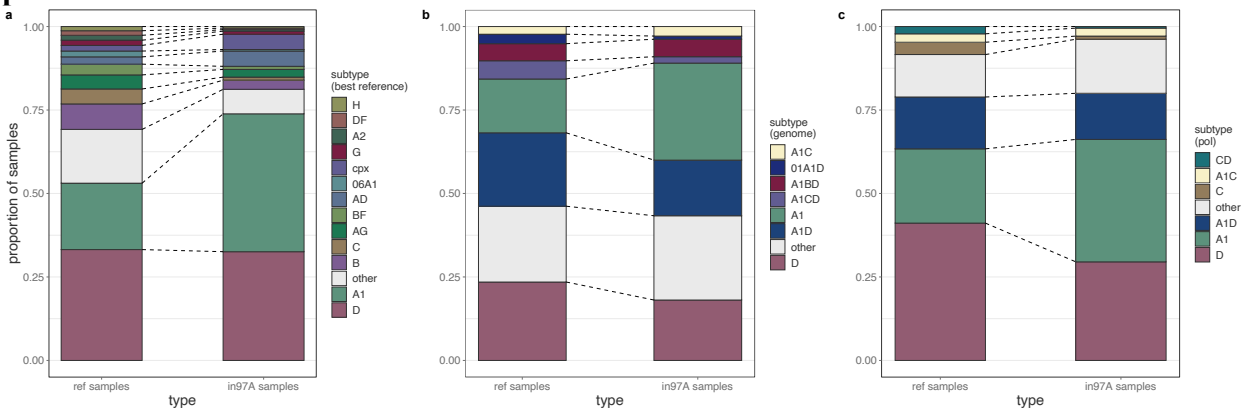
	R15 (2012)			R16 (2014)			R17 (2015)		
	Prevalence % (95% CI)	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**
Age category									
(14,24]	6.02 (2.99, 12.11)	ref	ref	9.45 (4.31, 20.69)	1.57 (0.549, 4.49)	0.4	11.51 (6.97, 19.01)	1.913 (0.809, 4.524)	1.40E-01
(24,34]	8.21 (5.13, 13.14)	1.364 (0.587, 3.17)	0.47	7.45 (3.84, 14.43)	0.578 (0.153, 2.181)	0.42	7.61 (5.21, 11.13)	0.485 (0.169, 1.388)	0.18
(34,50]	8.47 (4.83, 14.85)	1.407 (0.574, 3.451)	0.46	15.41 (7.48, 31.73)	1.159 (0.288, 4.669)	0.84	10.55 (7.13, 15.61)	0.652 (0.217, 1.957)	0.45
Community type									
Agrarian	7.66 (4.59, 12.78)	ref	ref	11.27 (5.48, 23.19)	1.472 (0.608, 3.565)	0.39	9.05 (5.73, 14.31)	1.182 (0.595, 2.349)	0.63
Fishing	6.38 (3.48, 11.69)	0.833 (0.377, 1.841)	0.65	7.75 (3.88, 15.49)	0.826 (0.231, 2.959)	0.77	9.11 (6.48, 12.81)	1.209 (0.455, 3.212)	0.7
Trading	11.62 (7.07, 19.1)	1.518 (0.744, 3.097)	0.25	10.82 (5.38, 21.79)	0.633 (0.184, 2.169)	0.47	10.25 (6.2, 16.94)	0.746 (0.278, 1.998)	0.56
Sex									
F	7.76 (4.87, 12.38)	ref	ref	11.45 (6.83, 19.18)	1.475 (0.736, 2.959)	0.27	11.08 (8.03, 15.28)	1.428 (0.81, 2.517)	0.22
M	7.9 (5.04, 12.39)	1.019 (0.533, 1.947)	0.96	8.98 (4.34, 18.56)	0.77 (0.256, 2.317)	0.64	7.76 (5.42, 11.12)	0.688 (0.307, 1.543)	0.36
Subtype									
A1	11.62 (6.84, 19.74)	ref	ref	21.3 (10.91, 41.6)	1.833 (0.781, 4.305)	0.16	13.16 (9.34, 18.55)	1.132 (0.602, 2.13)	7.00E-01
D	6.1 (3.19, 11.69)	0.525 (0.227, 1.215)	0.13	4.57 (1.87, 11.17)	0.408 (0.101, 1.65)	0.21	6.79 (4.52, 10.21)	0.982 (0.364, 2.653)	0.97
Other (recombinant)	8.49 (4.37, 16.5)	0.731 (0.312, 1.71)	0.47	11.25 (5.2, 24.34)	0.723 (0.191, 2.729)	0.63	9.75 (5.35, 17.79)	1.014 (0.339, 3.034)	0.98

	R18 (2017)			R19 (2019)		
	Prevalence % (95% CI)	**Prev. ratio (95% CI)**	***p*-value**	**Prevalence % (95% CI)**	**Prev. ratio (95% CI)**	***p*-value**
Age category						
(14,24]	9.26 (5.12, 16.72)	1.538 (0.615, 3.845)	0.36	40.73 (24.7, 67.16)	6.768 (2.864, 15.996)	<0.0001
(24,34]	10.42 (7.01, 15.48)	0.825 (0.274, 2.487)	0.73	18.94 (11.02, 32.55)	0.341 (0.111, 1.045)	0.06
(34,50]	11.26 (6.74, 18.81)	0.864 (0.263, 2.843)	0.81	10.53 (4.14, 26.75)	0.184 (0.046, 0.735)	0.017
Community type						
Agrarian	9.91 (5.91, 16.64)	1.294 (0.625, 2.681)	0.49	26.39 (15.89, 43.83)	3.446 (1.676, 7.083)	0.00077

Fishing	10.26 (6.77, 15.55)	1.243 (0.442, 3.497)	0.68	22.88 (13.24, 39.54)	1.042 (0.351, 3.095)	0.94
Trading	10.97 (6.42, 18.73)	0.729 (0.26, 2.044)	0.55	9.29 (3.07, 28.13)	0.232 (0.056, 0.952)	0.043
Sex						
F	8.46 (5.36, 13.35)	1.09 (0.568, 2.094)	0.8	18.62 (10.89, 31.84)	2.399 (1.178, 4.885)	0.016
M	11.97 (8.45, 16.95)	1.389 (0.584, 3.3)	0.46	23.7 (14.68, 38.26)	1.25 (0.475, 3.29)	0.65
Subtype						
A1	14.21 (9.46, 21.35)	1.223 (0.627, 2.386)	5.60E-01	27.99 (16.53, 47.38)	2.409 (1.141, 5.084)	0.021
D	8.57 (5.51, 13.35)	1.148 (0.409, 3.223)	0.79	17.41 (10.11, 29.99)	1.184 (0.383, 3.664)	0.77
Other (recombinant)	10.38 (5.14, 20.97)	1 (0.309, 3.239)	1	18.96 (5.49, 65.41)	0.927 (0.189, 4.552)	0.93

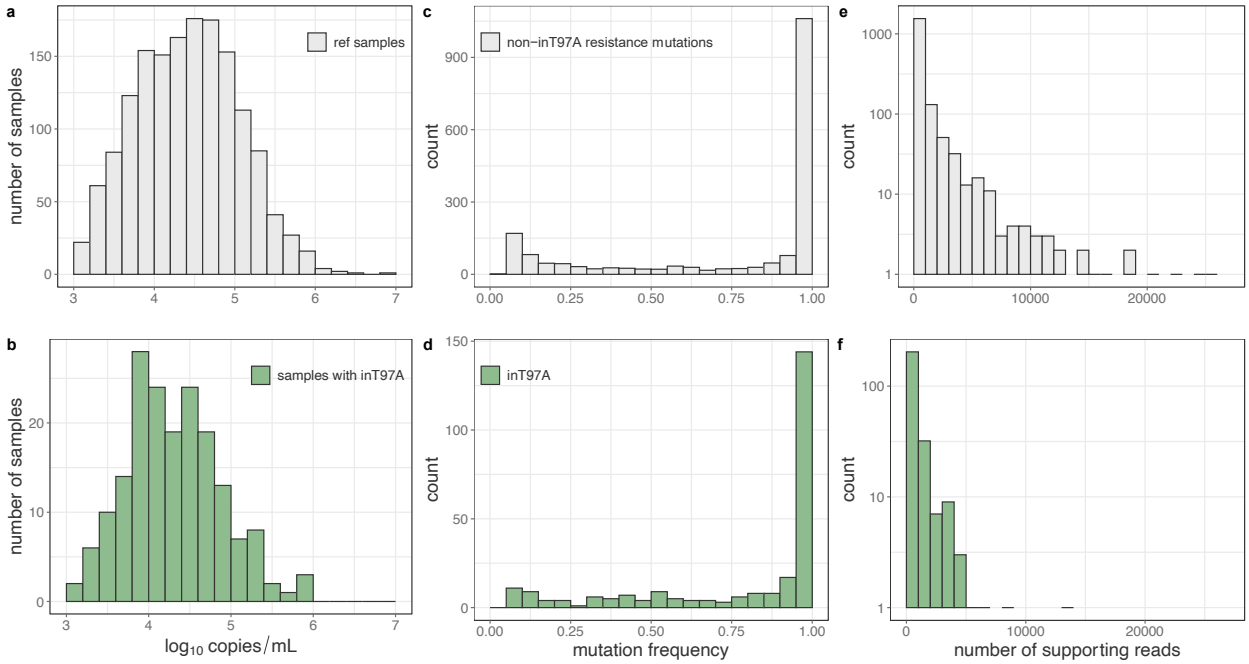
Estimates were generated using robust log-Poisson regression. 95% confidence intervals indicate the Wald confidence interval around the mean value in each category. Only the three most common subtypes are shown. Subtypes were inferred by identifying the closest reference match to each sequenced sample.

Supplementary Figure 7: Subtype of viremic pre-treatment PLHIV stratified by presence/absence of inT97A



Subtyping was done by identifying the closest reference match (A), using the whole genome (B), or using only the pol gene (C). Only subtypes observed in ≥ 3 sequences are shown. PLHIV = people living with HIV. In = integrase.

Supplementary Figure 8: Virological characteristics of viremic pre-treatment PLHIV stratified by presence/absence of inT97A



(A,B) Viral load of samples in which we do not (A) and do (B) identify inT97A. Samples with missing viral loads are removed. (C,D) Sequencing read frequency of identified non-inT97A (C) and inT97A (D) drug resistance substitutions. (E,F) Supporting sequencing read count of identified non-inT97A (E) and inT97A (F) drug resistance substitutions. In = integrase.

Supplementary References

1 Bonsall D, Golubchik T, de Cesare M, *et al.* A Comprehensive Genomics Solution for HIV Surveillance and Clinical Monitoring in Low-Income Settings. *J Clin Microbiol* 2020; **58**: e00382-20.

2 Halekoh U, Højsgaard S, Yan J. The *R* Package **geepack** for Generalized Estimating Equations. *J Stat Softw* 2006; **15**. DOI:10.18637/jss.v015.i02.

3 Wohl S, Lee EC, DiPrete BL, Lessler J. Sample Size Calculations for Variant Surveillance in the Presence of Biological and Systematic Biases. 2022; published online Jan 1. DOI:10.1101/2021.12.30.21268453.