OXFORD UNIVERSITY PRESS | Journal of Antimicrobial Chemotherapy

Surveillance of HIV-1 transmitted integrase strand transfer inhibitor resistance in the UK

Journal:	*Journal of Antimicrobial Chemotherapy*
Manuscript ID	JAC-2020-0476.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Mbisa, Jean; Public Health England Colindale, Virus Reference Department Ledesma, Juan; Public Health England, Virus Reference Department Kirwan, Peter; Public Health England, HIV and STI Department Bibby, David; Public Health England Colindale, Virus Reference Department Manso, Carmen; Public Health England, Virus Reference Department Skingsley, Andrew; Public Health England, HIV and STI Department Murphy, Gary; Public Health England Colindale, Virus Reference Department Brown, Alison; Public Health England, HIV and STI Department Dunn, David; University College London, Institute of Global Health Delpech, Valerie; Public Health England Colindale, HIV and STI Department Geretti, Anna; Institute of Infection and Global Health (IGH), Department of Clinical Infection, Microbiology and Immunology (CIMI); Anna Maria Geretti,
Keywords:	Integrase strand transfer inhibitor, Recent infection, HIV, transmitted drug resistance

Surveillance of HIV-1 transmitted integrase strand transfer inhibitor resistance in the UK ~~and implications on clinical practice~~

Jean L. MBISA[1,2,*†], Juan LEDESMA[1,2†], Peter KIRWAN[1], David F. BIBBY[1], Carmen MANSO[1], Andrew SKINGSLEY[1], Gary MURPHY[1], Alison BROWN[1], David T. DUNN[3], Valerie DELPECH[1,2] and Anna Maria GERETTI[4]

[1] National Infection Service, Public Health England, London UK

[2] National Institute for Health Research (NIHR) Health Protection Research Unit in Blood Borne and Sexually Transmitted Infections, London, UK

[3] Institute for Global Health, University College London, London UK

[4] Institute of Infection and Global Health, University of Liverpool, Liverpool, UK.

Running Title. No transmitted InSTI resistance in UK

Summary. Integrase strand transfer inhibitor (InSTI) drug resistance data from HIV-1 recently infected individuals in the UK shows no evidence of major InSTI resistance-associated mutations as a high-frequency variant. These data suggest no clinical benefit for baseline integrase resistance testing.

* Corresponding author: Antiviral Unit, Virus Reference Department, National Infection Service, Public Health England, 61 Colindale Avenue, London NW9 5EQ, UK. (tamyo.mbisa@phe.gov.uk)

† These authors contributed equally to this work

Abstract

Background: HIV treatment guidelines have traditionally recommended that all HIV-positive individuals are tested for evidence of drug resistance prior to starting ART. Testing for resistance to reverse transcriptase inhibitors (RTI) and protease inhibitors (PI) is well established in routine care. However, testing for integrase strand transfer inhibitors (InSTI) resistance is less consistent.

Objectives: To inform treatment guidelines by determining the prevalence of InSTI resistance in a national cohort of recently infected individuals.

Patients and methods: Recent (within 4 months) HIV-1 infections were identified using a Recent Infection Testing Algorithm of new HIV-1 diagnoses in the UK. Resistance-associated mutations (RAMs) in integrase, protease and RT were detected by ultradeep sequencing which allows for the sensitive estimation of the frequency of each resistant variant in a sample.

Results: The analysis included 655 randomly selected individuals (median age 33 years, 95% male, 83% men who have sex with men, 78% white) sampled in 2014 to 2016 and determined to have a recent infection. These comprised 320, 138 and 197 samples from 2014, 2015 and 2016, respectively. None of the samples had major InSTI RAMs occurring at high variant frequency (≥20%). A subset (25/640, 3.9%) had major InSTI RAMs occurring only as low-frequency variants (2-20%). In contrast, 47/588 (8.0%) had major RTI and PI RAMs at high frequency.

Conclusions: Between 2014 and 2016, major InSTI RAMs were uncommon in adults with recent HIV-1 infection, only occurring as low-frequency variants of doubtful clinical significance. Continued surveillance of~~Testing~~ newly diagnosed patients for evidence of transmitted InSTI resistance is recommended to inform~~unlikely to be of~~ clinical practice~~benefit at present~~.

Introduction

In 2007, raltegravir was the first integrase strand transfer inhibitor (InSTI) introduced into clinical practice, initially for treatment-experienced HIV-positive patients requiring rescue therapy[1] and two years later for all patients including those who are treatment-naïve. RAL was followed by elvitegravir in 2012 as part of a fixed dose, single tablet combination (Stribild) that includes the booster cobicistat, emtricitabine and tenofovir disoproxil fumarate. Elvitegravir was subsequently reformulated in combination with tenofovir alafenamide instead of tenofovir disoproxil fumarate (Genvoya). Second-generation InSTIs comprise dolutegravir, which was approved in 2013, and more recently bictegravir coformulated with tenofovir alafenamide and emtricitabine (Biktarvy), which was approved in 2018.[2] Large clinical trials have demonstrated that the second-generation InSTIs are potent suppressors of HIV replication and have good safety and high genetic barriers to the emergence of drug resistance.[3-6] These features make them the preferred third agent for starting antiretroviral therapy (ART) in combination with a backbone of two nucleoside reverse transcriptase inhibitors (NRTIs).[7 8] First-line regimens based on non-nuclease reverse transcriptase inhibitors (NNRTIs) or boosted protease inhibitors (bPIs) are instead reserved for selected scenarios.

Mirroring European and American guidelines, the British HIV Association (BHIVA) guidelines for the treatment of HIV-1 positive adults recommend that resistance testing by viral partial genome sequencing be undertaken in all newly diagnosed patients prior to starting ART to allow the detection of transmitted drug resistance (TDR).[7-10] Sequencing should be performed for reverse transcriptase and protease genes, based on studies showing that when the prevalence of TDR in the population exceeds a threshold of 1-5% it is cost-effective to screen patients to guide treatment selection.[11 12] In the UK, the prevalence of TDR affecting NRTIs, NNRTIs or PIs peaked at ~14% in 2002 and has remained stable at 7-9% since 2006.[13 14] To date, there is no recommendation for baseline integrase sequencing as little evidence exists of the transmission of InSTI resistance-associated mutations (RAMs) in the UK and worldwide.[15]

Most routine resistance testing is performed using conventional Sanger sequencing technology, which has variant frequency detection threshold of ~20% and hence fails to detect variants that are present below this threshold in a patient's viral population. Next generation sequencing (NGS) technologies allow the detection of variants present in a sample at frequency as low as 1%.[16] The clinical significance of low-frequency resistant variants remains under debate. It has been shown that low-frequency variants with mutations affecting the NNRTIs, and to a lesser extent the NRTIs, significantly reduce responses to first-line therapy with 2 NRTIs plus 1 NNRTI, while showing no appreciable effect on bPI-based regimens.[17-20] However, transmission is unlikely to be a source of the large majority of these low-frequency variants in individuals who have recently acquired HIV and consequently would have minimal to no impact on treatment outcome as they would not have been selected under drug pressure.[21]

Several studies have reported no evidence of InSTI major RAMs in treatment-naïve or recently infected HIV-1 positive populations using Sanger sequencing.[22-27] The few studies reporting apparent transmission of InSTI RAMs included mutations that are polymorphic among ART-naïve patients (e.g. L74IM, T97A and E157Q).[28 29] One notable exception is a study from Taiwan that observed InSTI major RAMs (e.g. Q148HKR and Y143R) in 1.2% of 1307 ART-

naïve individuals under a specific epidemiological circumstance where there was a large reservoir of InSTI resistance among the treated population.[30]

The UK national reference laboratory receives blood samples from half of all newly diagnosed cases of HIV-1 infection for incidence testing using a Recent Infection Testing Algorithm (RITA).[31] Recently infected individuals are the most relevant population as they are treatment-naïve and detection of resistance is most likely due to transmission and prior to natural decay. Using this resource, we determined the national prevalence of TDR to InSTIs, NRTI, NNRTI and bPI by performing NGS on samples from newly diagnosed patients identified as infected in the previous 4 months. The findings will inform the clinical utility of baseline resistance testing for integrase in the UK.

Patients and methods

Study population

The UK national reference laboratory applies ~~recent infection testing algorithm (~~RITA~~)~~ to new HIV-1 diagnoses as a sentinel national surveillance programme. Blood samples from ART-naïve individuals collected at HIV or Genitourinary Medicine Clinics in the UK are identified as likely recent infections (within 4 months of sample collection) using a Limiting-antigen (LAg) avidity assay with an OD index <1.5. The assay differentiates likely recent from long standing infection by the strength of HIV-specific antibody-antigen binding.[32] The assay has a misclassification rate of long-standing HIV infections as recent of <1% when RITA is applied and samples close to the OD index cut-off values are more likely to be misclassified.[33] The RITA algorithm also includes matching the sample to individual HIV records of the HIV and AIDS Reporting System (HARS). Individuals with an OD index <1.5 must also have a CD4+ cell count (>200 cells/mm^3) and viral load (>1,000 copies/mL) to be assigned as 'recent infections'. In 2014-2016, RITA was applied to 8,379 (47.9%) of new diagnoses and 1,765 (21.1%) were identified as recent infections (**Figure 1**). We randomly selected 655 (37.1%) of these plasma specimens with residual volume for NGS analysis, comprising by year 320, 138 and 197 samples collected in 2014, 2015 and 2016, respectively. Linked demographic and clinical information was extracted from HARS. Prescription data were used to determine the use of InSTIs in clinical practice in England between 2010 and 2015.

Next generation sequencing

~~RNA was extracted using QIAsymphony virus/pathogen DSP mini kit (Qiagen) and eluted in a final volume of 60 µl.~~ Samples collected in 2014 were extracted using QIAsymphony virus/pathogen DSP mini kit (Qiagen) using 200 µl of plasma, eluted in a final volume of 60µL and processed using a previously described PCR amplicon-based NGS assay for protease-RT[34] and integrase was amplified in a nested PCR reaction. Briefly, cDNA was generated using 20 µl of RNA, Qiagen OneStep RT-PCR Kit (Qiagen) and primers H10F2 (5'-GCACAYAARGGRATTGGAGGAAATGA-3') and H10R3 (5'-CCTAGTGGRATGTGTACTTCTGA-3'), both at 1~~5~~ µM under the following cycling conditions: 50˚C for 40 minutes, 95 ˚C for 15 minutes, 35 cycles of 95˚C for 30 seconds, 53˚C for 30 seconds and 72˚C for 1 minutes with a final elongation step at 72˚C for 4 minutes. Two µl of cDNA were then used in a semi-nested PCR using Platinum Taq DNA Polymerase kit

(Invitrogen) and ~~10~~0.4 µM of primer H10F2 and ~~40~~1.6 µM of primer H10R2 (5'-CATATGRTGYTTTACTAAACTHTTCCA-3') under the following cycling conditions: 95˚C for 5 minutes, 35 cycles at 94˚C for 30 seconds, 53˚C for 30 seconds and 72˚C for 1 minutes and final elongation step at 72˚C for 2 minutes. Amplicons for protease-RT and integrase were pooled in equimolar concentration and then sequenced as previously described.[34]

Samples collected in 2015 and 2016 were processed using a sequence capture whole genome sequencing (WGS) assay ~~(see supplementary files for details)~~. Briefly, 350µl of plasma was extracted using the NucliSENS system on the easyMag platform (bioMérieux) and eluted into a volume of 25µl, all of which was subjected to DNAse digestion with 0.25U of TURBO DNase (Thermo Fisher Scientific) in 30µl reactions incubated for 30 minutes at 37°C. Digestion products were cleaned-up using 2X AMPure XP Beads (Beckman Coulter) following the manufacturer's instructions, with a final elution volume of 10µl nuclease-free water. The 10µl volume of DNAse-digested RNA extracts were used to generate DNA libraries, using the KAPA RNA HyperPrep Kit (Roche). DNA libraries were pooled in a total of 500ng and hybridized using 120-nt HIV-specific biotinylated oligonucleotide probes and NimbleGen SeqCap target enrichment reagents (Roche) following the manufacturer's specifications. Following hybridization, the HIV DNA libraries bound to the biotinylated probes were partitioned using magnetic streptavidin-coated beads and subjected to a further 14 cycles of PCR amplification. The concentration of the final pool was quantified using the KAPA SYBR FAST Universal qPCR Kit for Illumina libraries (KAPA Biosystems) on a 7500 Real-Time PCR System (Applied Biosystems), and analysed for fragment size distribution using the High Sensitivity DNA Kit (Agilent) on a 2100 Bioanalyser Instrument, following both manufacturers' specifications. Sequencing was performed on an Illumina MiSeq instrument using the MiSeq Reagent Kit V2 (300 cycles) (Illumina) according to the manufacturer's guidelines, with the following minor modifications. The final pool was diluted to 2nM and denatured with 0.2N sodium hydroxide for 2 minutes, incubated for 4 minutes at 95°C, and diluted in kit reagent HT1 to produce 1ml of a 20pM solution. These were further diluted to make 700µl of a 9pM solution, of which 10% was substituted with 12.5pM PhiX (Illumina). A total of 600µl of this final solution were loaded onto the MiSeq cartridge.

Bioinformatic analysis

MiSeq paired end FASTQs were trimmed for quality with trimmomatic (v0.39, with LEADING and TRAILING set to 30 and MINLEN to 50)[35], and human reads removed through BWA (v0.7.17) mapping to the human genome (GRCh37), retaining unmapped pairs. Contigs generated from dehumanised reads by *de novo* assembly using SPAdes (v3.13.1)[36] were split into fragments of approximately 500nt (depending on length) and BLASTed against 2,427 reference genome sequences annotated and aligned by LANL (http://www.hiv.lanl.gov/). An in-house Python script was used to build a draft sequence from the contigs and their locations within the genome alignment, filling gaps from the reference sequence(s) informing the flanking termini (in a fashion somewhat analogous to LASTZ). Two rounds of BWA mapping and consensus derivation (using an in-house C++ script – QuasiBAM)[37] were performed to obtain the final nucleotide frequency table, from which an in-house perl script derived consensus genome sequences at 20% and 2% nucleotide frequencies at a minimum read depth of 30, the latter frequency previously established as the minimum threshold of the assay.[34]

The positive percent agreement (PPA) between amplicon and sequence capture methods was 99.8% [99.7-99.9; 95%CI] and 99.0% [98.8-99.2] at 20% and 2% variant frequency thresholds, respectively at the nucleotide level (**Tables S1 and S2, Figure S1**).

Drug resistance analysis

The generated consensus sequences were analysed for surveillance drug resistance mutations (SDRMs) using the Calibrated Population Resistance (CPR) tool that uses the WHO 2009 list of SDRMs for PI and RTI, and recently proposed InSTI SDRMs list 2019: T66AIK, E92GQ, G118R, F121Y, E138AKT, G140ACS, Y143CHRS, S147G, Q148HRK, N155H, S230R and R263K.[38-40] The integrase sequences were also analysed for the presence of the following InSTI accessory mutations: H51Y, Q95K, T97A, A128T, V151A, S153FY, E157Q and G163KR. These mutations have minimal, if any, effect on InSTI susceptibility when present alone but may contribute to reduced susceptibility in combination with InSTI major RAMs. For determination of mutational load, viral load and CD4+ count data was only used if performed within 30 days of the sample used for RITA and sequencing. Additional mutational load data for PI, NRTI and NNRTI was obtained from previously reported RT and protease sequencing of recently infected individuals from 2011 to 2013.[34]

Statistical analyses

Descriptive statistics (median and interquartile ranges) are provided for continuous variables whereas frequency distributions are provided for categorical variables. Mutational load datasets were compared using Mann-Whitney U-test with significance level set at $p<0.05$.

Sequence data

Consensus HIV-1 pol sequences from this study have been submitted to GenBank and may be accessed by the following accession numbers: MT570368-MT571329.

Results

~~Use of InSTIs in England~~

~~The proportion of adults using InSTIs increased significantly from a low of 2.4% in London in 2010 to between 17.4% and 25.7% in the Midlands and East of England and the North of England, respectively in 2015 (**Figure 2**).~~

Study population characteristics

The proportion of adults using InSTIs increased significantly from a low of 2.4% in London in 2010 to between 17.4% and 25.7% in the Midlands and East of England and the North of England, respectively in 2015 (**Figure 2**).

Sequencing was performed on plasma samples from 655 recently infected individuals collected between 2014 and 2016. The characteristics of the study population are summarised in **Table 1**. The majority were male (94.5%) and of white ethnic background (77.9%) and their risk factor for HIV infection was classed as having sex with other men (82.9%). Most of the recent infections were from the London region (57.1%) and the median age of the study population was 33 years [26.5-41; IQR]. Median viral load was log 5.18 [4.61-6.03] copies/mL and median CD4+ count

was 545 [408-723] cells/μL, in keeping with recency of infection. Subtyping using the *pol* gene showed most individuals were infected with subtype B (67.5%).

Prevalence of integrase resistance-associated mutations

Of the 655 samples sequenced, 640 (97.7%) generated good quality integrase gene sequence data (complete gene coverage at minimum read depth of 100); 316, 132 and 192 in 2014, 2015 and 2016, respectively. No InSTI major RAMs were detected in the 640 sequences as a high-frequency variant (≥20%). A total of 25 (3.9%) sequences contained major InSTI ~~major~~ RAMs as low-frequency variants occurring at a frequency between 2% and 20% (**Figure 3A**). By year, 18 (5.7%), 3 (2.3%) and 4 (2.1%) sequences contained InSTI major RAMs as low-frequency variants in 2014, 2015 and 2016, respectively. In contrast, 39 (6.1%) sequences contained InSTI accessory mutations as a high-frequency variant and 18 (2.8%) as a low-frequency variant (**Figure 3A**). By year, the number of sequences containing InSTI accessory mutations were 19 (6.0%), 11 (8.3%) and 9 (4.7%) as a high-frequency variant, and 8 (2.5%), 7 (5.3%) and 3 (1.6%) as a low-frequency variant in 2014, 2015 and 2016, respectively.

The InSTI major RAMs and accessory mutations detected are shown in **Figure 3B**. The most common InSTI major RAMs detected as low-frequency variants were E138K (8/25; 32.0%) and Q148K (7/25; 28.0%). On the other hand, the most common InSTI accessory mutation detected as a high-frequency variant was E157Q (29/39; 74.4%) whereas V151A was the most common InSTI accessory mutation detected as a low-frequency variant (5/18; 27.8%). Most of the InSTI accessory mutations present as high-frequency variants were associated with non-B subtypes (22/39; 56.4%) with the majority associated with the circulating recombinant forms CRF02_AG (n=9) and CRF06_cpx (n=8). The presence of low frequency RAMs was confirmed by read-based RAM analysis using the variant frequency files produced by the QuasiBam software (**Table S3**).

Prevalence of ~~R~~reverse ~~T~~transcriptase and ~~P~~protease ~~R~~resistance-associated mutations

Of the 655 samples sequenced, 619, 593 and 588 generated good quality sequence data for protease, reverse transcriptase and both gene regions, respectively. By year, 302, 129 and 188 protease, and 295, 116 and 182 RT sequences were generated for 2014, 2015 and 2016, respectively. TDR prevalence for all drug classes for the period 2014-2016 was 8.0% (47/588) for high-frequency variants and 10.9% (64/588) for low-frequency variants. TDR mutations were detected as a high-frequency variant in 15 (2.4%), 20 (3.4%) and 17 (2.9%) sequences against PIs, NRTIs and NNRTIs, respectively (**Figure 4A**). The overall prevalence of TDR low-frequency variants against PIs, NRTIs and NNRTIs was 27 (4.4%), 29 (4.9%) and 12 (1.0%), respectively (**Figure 4A, Table S3**). The most common TDR mutations detected as a high-frequency variant were L90M (7/15; 46.7%), T215rev (11/20; 55.0%) and K103N (12/17; 70.6%) against PIs, NRTIs and NNRTIs, respectively (**Figure 4B**). In contrast, the most common TDR mutations detected as a low-frequency variant were M46IL (13/27; 48.1%), D67GNE (13/29; 44.8%) and G190E (6/12; 50.0%) against PIs, NRTIs and NNRTIs, respectively (**Figure 4B**).

Of the 655 sequences 581 (88.7%) generated sufficient and good quality data in all three polymerase gene regions. Two (0.3%) of the samples had a low-frequency ~~variant~~ InSTI major

RAM and a high-frequency ~~variant~~ RAM in protease or RT: Q148K + G190A (NNRTI) and E92G + T215S (NRTI). Both were subtype B. Three samples (0.5%) had a low-frequency ~~variant~~ InSTI major RAM (E138K) and low-frequency ~~variant~~ RAM in protease and/or RT: D67N (NRTI) + M46I (PI), M46I (PI) and D67G (NRTI) and they were subtype F, B and CRF01_AE, respectively.

Mutational load of low frequency ~~RAMs~~resistance-associated mutations

We determined the mutational load of low-frequency RAMs as previously described.[19]~~.~~ The median mutational load of low-frequency InSTI RAMs was 3,833 [895-14,733; IQR] copies/mL (**Figure 5**). The median mutational load of low-frequency InSTI RAMs was similar to that of low-frequency PI and NRTI RAMs at 5,914 [1,470-31,552] and 2,706 [1,387-11,726] copies/mL, respectively. In contrast, the mutational load for low-frequency NNRTI RAMs was slightly higher and had a broad range at 10,188 [2,170-95,313] copies/mL; however, the difference was~~is~~ not statistically significant (p>0.05, Mann Whitney U test).

Discussion

Surveillance of transmitted InSTI resistance among ~~a population of~~ 655 ~~recently infected~~ individuals with recent HIV infection in the UK who were sampled between 2014 and 2016 showed no evidence of ~~transmission of~~major InSTI ~~major~~ RAMs when considering mutants present at high frequency in the individuals' samples. In contrast, at approximately 8%, the prevalence of TDR to PIs and RTIs remains steady compared to last reported figures in 2014 ~~at approximately 8%~~. This prevalence is still higher than the recommended threshold of 1-5% where baseline resistance testing is considered of benefit at the population level.

On the other hand, accessory InSTI resistance mutations were detected as ~~a~~ high-frequency variants in 6.1% of the study population. The most common accessory InSTI RAMs were T97A and E157Q; both ~~which~~ are polymorphic and observed at a high prevalence (up to 7%) in InSTI-naïve individuals infected with non-B subtypes e.g. CRF02_AG~~, at up to 7% prevalence~~. They are also selected in patients experiencing treatment failure with first-generation InSTIs, raltegravir and elvitegravir; however, they have little effect on InSTI susceptibility when present alone.[41]

Ultradeep sequencing allowed the detection of mutations below the Sanger sequencing variant frequency threshold of ~20%. Major InSTI ~~major~~ RAMs were detected as a low-frequency variant in 3.9% of the study population, at a variant frequency between 2% and 20%. It has been argued that the mutational load of low-frequency RAMs, especially for NNRTIs, could play a role in treatment failure.[42] The mutational load of low-frequency InSTI RAMs was comparable to that of NRTI and PI RAMs, but was lower and had a very narrow range compared to that of NNRTI RAMs. Compared to the data for low-frequency NNRTI RAMs, ~~T~~there is less compelling evidence that low-frequency PI and NRTI RAMs contribute to treatment failure.[20] It is likely that the mutational load of low-frequency RAMs is associated with the impact on virus replication fitness, therefore InSTI, NRTI and PI RAMs which have a high impact on virus replication fitness are unlikely to accumulate to high absolute levels compared to NNRTI RAMs. In addition, ~~W~~we recently showed evidence that the majority of low-frequency RAMs to PIs and RTIs in recently infected individuals are not a result of a transmission event and thus would not have been

selected under drug pressure.[21] Furthermore, a recent study showed no association between the presence of low-frequency InSTI RAMs prior to initiation of treatment and treatment outcomes.[43] Taken together, these data suggest that the low-frequency InSTI RAMs in recently infected individuals are less likely to affect treatment outcome especially as current second-generation InSTIs, dolutegravir and bictegravir, are highly effective and have very high ~~genetic~~ barriers to resistance. However, ~~the~~ low-frequency InSTI RAMs may still have an impact in ~~treatment-experienced~~ individuals~~, those~~ with poor adherence or those who harbour resistance to other components of their ART regimen~~who have undergone treatment interruption~~.

The proportion of individuals on an ART regimen that included an InSTI was approximately 20% during the period covered by the study, as estimated using prescription data from NHS England. The use of InSTIs as part of first-line regimens is anticipated to continue to rise in the UK reflecting national and international treatment guidelines.[44] The use of InSTIs in the UK increased from less than 10% in 2014 to over 20% in 2015; thus, the effect of this and further projected increases in InSTI use may not be captured in this surveillance study. Nonetheless, the virological suppression rate for people on InSTI-based therapy in the UK is very high (>95%) and the likelihood of the emergence of drug resistance for those failing dolutegravir or bictegravir plus two NRTIs in first-line ART is negligible.[3 4] However, raltegravir and elvitegravir have been used for longer than dolutegravir and bictegravir in ART-naïve and ART-experienced patients and these drugs are more likely to result in treatment failure with resistance selection. Thus, these groups may have generated a pool of potential transmitters which may later contribute to transmitted InSTI resistance. All these factors necessitate continued surveillance of InSTI TDR in the coming years.

Reflecting the focus on recent infection, ~~A~~another limitation of the study is that most of the sampled population was ~~from England, were~~ male~~,~~ of white ethnic background from England~~and~~ whose probable route of HIV exposure was sex between men and who were infected with subtype B virus. This is because gay and bisexual men are more likely to have recently acquired infection at HIV diagnosis. Thus, these findings may not be generalizable to the whole of the UK population living with HIV, and particularly women and those infected with non-B subtypes. In addition, the frequency of InSTI use may be different in Scotland, Wales and Northern Ireland, data which was not captured in this study.

Two different sequencing methods were used in this study, an amplicon-based approach using gene-specific nested PCR followed by DNA library prep and a sequence capture approach which is dependent on RNA library prep followed by enrichment using HIV-1 specific probe baits. Overall the consensus sequence generated by both methods were highly concordant (>99%) using the 20% and 2% variant frequency at nucleotide level (see supplementary data). Discordances were at mixed base positions where one method detected only one of the mixed bases with a tendency for low-frequency variants detected by amplicon method being infrequently detected by sequence capture whereas the opposite was true. This could either be due an overcall of low-frequency variants by the amplicon method or a decreased sensitivity for detection of low-frequency variants by the sequence capture. This requires further investigation using standardised reference or control material with well-characterised low-frequency variants

at specific frequencies. The validation of NGS methods to accurately reflect *in vivo* low-frequency variants is essential to determine their effect on clinical outcomes.

In conclusion, this study shows no evidence of transmitted InSTI resistance in the recently infected population in the UK. However, performing baseline integrase resistance testing is still important, especially for national reference laboratories in order to provide surveillance data ~~suggesting that baseline InSTI resistance testing is unlikely to be of clinical benefit. However, continued surveillance is recommended~~ and possibly in selected patients in routine clinical practice. One consideration is~~considering~~ that the use of InSTIs as part of first-line cART is anticipated to continue to increase worldwide following WHO recommendations. The large-scale use of DTG in resource-limited settings is likely to take place with limited viral load monitoring and thus could result in significant increases in transmitted InSTI resistance from a more global perspective~~in the UK and low-frequency InSTI resistance was detected at 3.9%~~. In parallel, prospective cohort studies to assess treatment outcomes in recently infected individuals harbouring ~~the~~ low-frequency RAMs~~variants~~ would best inform their clinical significance and diagnostic utility. The use of WGS adopted from 2015 onwards will also be useful in analysing other regions of the HIV-1 genome that have been postulated to be involved in development of resistance to second-generation InSTI, such as envelope and 3' polypurine tract (PPT).[45 46]

Funding

This work was supported by ViiV Healthcare; Public Health England (PHE); and by the National Institute for Health Research and undertaken by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Blood Borne and Sexually Transmitted Infections at University College London (UCL) in partnership with PHE, in collaboration with the London School of Hygiene and Tropical Medicine. The views expressed are those of the authors and not necessarily those of the NIHR, the Department of Health and Social Care or PHE.

Conflict of interest

~~None~~AMG reports consultancy fees and research funding paid to the University of Liverpool from Gilead Sciences, Janssen, and ViiV Healthcare, and personal fees from Roche Pharma Research & Early Development, outside of the presented work.

Acknowledgements

We thank staff of Antiviral Unit and Clinical Service Unit in the Virus Reference Department, PHE for providing laboratory support. We also wish to thank the NIHR HPRU in Blood Borne and Sexually Transmitted Infections Steering Committee: Caroline Sabin (Director), John Saunders (PHE Lead), Catherine Mercer, Gwenda Hughes, Jackie Cassell, Greta Rait, Samreen Ijaz, Tim Rhodes, Sema Mandal, Kholoud Porter, William Rosenberg for reviewing the manuscript.

Author Contributions

JLM and JL contributed equally to this manuscript. JLM, AMG, VD and DTD conceived the hypotheses and designed the study. JL and CM performed the sequencing experiments. all the laboratory work. PK, AS, GM, AB and VD collected the metadata and coordinated RITA testing and HARS database. JLM, JL and DFB performed bioinformatics analyses. JLM and AMG drafted the manuscript. All authors provided critical reading that shaped the manuscript and approved the final version of the manuscript. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted and had final responsibility for the decision to submit for publication.

Ethics statement

PHE has Section 251 approval, which is reviewed annually, and provides the legal basis for the collection of HIV patient-level data for public health monitoring purposes. In addition, the HIV surveillance dataset is reviewed annually by the PHE Caldicott Panel to ensure compliance with information governance policies.

References

1. Arribas JR, Eron J. Advances in antiretroviral therapy. *Curr Opin HIV AIDS* 2013;**8**(4):341-9. doi: 10.1097/COH.0b013e328361fabd [published Online First: 2013/05/15]
2. Mesplede T, Wainberg MA. Integrase Strand Transfer Inhibitors in HIV Therapy. *Infect Dis Ther* 2013;**2**(2):83-93. doi: 10.1007/s40121-013-0020-8 [published Online First: 2014/08/20]
3. Walmsley S, Baumgarten A, Berenguer J, *et al*. Brief Report: Dolutegravir Plus Abacavir/Lamivudine for the Treatment of HIV-1 Infection in Antiretroviral Therapy-Naive Patients: Week 96 and Week 144 Results From the SINGLE Randomized Clinical Trial. *J Acquir Immune Defic Syndr* 2015;**70**(5):515-9. doi: 10.1097/QAI.0000000000000790 [published Online First: 2015/08/12]
4. Sax PE, Pozniak A, Montes ML, *et al*. Coformulated bictegravir, emtricitabine, and tenofovir alafenamide versus dolutegravir with emtricitabine and tenofovir alafenamide, for initial treatment of HIV-1 infection (GS-US-380-1490): a randomised, double-blind, multicentre, phase 3, non-inferiority trial. *Lancet* 2017;**390**(10107):2073-82. doi: 10.1016/S0140-6736(17)32340-1 [published Online First: 2017/09/05]
5. Gallant J, Lazzarin A, Mills A, *et al*. Bictegravir, emtricitabine, and tenofovir alafenamide versus dolutegravir, abacavir, and lamivudine for initial treatment of HIV-1 infection (GS-US-380-1489): a double-blind, multicentre, phase 3, randomised controlled non-inferiority trial. *Lancet* 2017;**390**(10107):2063-72. doi: 10.1016/S0140-6736(17)32299-7 [published Online First: 2017/09/05]
6. Acosta RK, Willkom M, Martin R, *et al*. Resistance Analysis of Bictegravir-Emtricitabine-Tenofovir Alafenamide in HIV-1 Treatment-Naive Patients through 48 Weeks. *Antimicrob Agents Chemother* 2019;**63**(5) doi: 10.1128/AAC.02533-18 [published Online First: 2019/02/26]
7. Churchill D, Waters L, Ahmed N, *et al*. British HIV Association guidelines for the treatment of HIV-1-positive adults with antiretroviral therapy 2015. *HIV Med* 2016;**17** Suppl 4:s2-s104. doi: 10.1111/hiv.12426 [published Online First: 2016/08/30]
8. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretrovirals agents in adults and adolescents with HIV. Department of Health and Human Services. Available at http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf. (Accessed 04 November 2019).
9. European AIDS Clinical Society. Guidelines. 2018. Available at https://www.eacsociety.org/files/2018_guidelines-9.1-english.pdf (Accessed 04 November 2019).
10. Gunthard HF, Calvez V, Paredes R, *et al*. Human Immunodeficiency Virus Drug Resistance: 2018 Recommendations of the International Antiviral Society-USA Panel. *Clin Infect Dis* 2019;**68**(2):177-87. doi: 10.1093/cid/ciy463 [published Online First: 2018/07/28]
11. Sax PE, Islam R, Walensky RP, *et al*. Should resistance testing be performed for treatment-naive HIV-infected patients? A cost-effectiveness analysis. *Clin Infect Dis* 2005;**41**(9):1316-23. doi: 10.1086/496984 [published Online First: 2005/10/06]
12. Hirsch MS, Brun-Vezinet F, Clotet B, *et al*. Antiretroviral drug resistance testing in adults infected with human immunodeficiency virus type 1: 2003 recommendations of an International AIDS Society-USA Panel. *Clin Infect Dis* 2003;**37**(1):113-28. doi: 10.1086/375597 [published Online First: 2003/06/28]
13. UK Collaborative Group on HIV Drug Resistance, UK Collaborative HIV Cohort Study, UK Register of HIV Seroconverters. Evidence of a decline in transmitted HIV-1 drug resistance in the United Kingdom. *AIDS* 2007;**21**(8):1035-9. doi: 10.1097/QAD.0b013e3280b07761 [published Online First: 2007/04/26]

14. Tostevin A, White E, Dunn D, *et al*. Recent trends and patterns in HIV-1 transmitted drug resistance in the United Kingdom. *HIV Med* 2017;**18**(3):204-13. doi: 10.1111/hiv.12414 [published Online First: 2016/08/02]
15. Geretti AM, Armenia D, Ceccherini-Silberstein F. Emerging patterns and implications of HIV-1 integrase inhibitor resistance. *Curr Opin Infect Dis* 2012;**25**(6):677-86. doi: 10.1097/QCO.0b013e32835a1de7 [published Online First: 2012/10/23]
16. Gega AK, M.J. New Technology to detect low-level drug-resistant HIV variants. *Future Virology* 2011;**6**:17-26.
17. Geretti AM, Paredes R, Kozal MJ. Transmission of HIV drug resistance: lessons from sensitive screening assays. *Curr Opin Infect Dis* 2015;**28**(1):23-30. doi: 10.1097/QCO.0000000000000136 [published Online First: 2014/12/17]
18. Li JZ, Paredes R, Ribaudo HJ, *et al*. Low-frequency HIV-1 drug resistance mutations and risk of NNRTI-based antiretroviral treatment failure: a systematic review and pooled analysis. *JAMA* 2011;**305**(13):1327-35. doi: 10.1001/jama.2011.375 [published Online First: 2011/04/07]
19. Cozzi-Lepri A, Noguera-Julian M, Di Giallonardo F, *et al*. Low-frequency drug-resistant HIV-1 and risk of virological failure to first-line NNRTI-based ART: a multicohort European case-control study using centralized ultrasensitive 454 pyrosequencing. *J Antimicrob Chemother* 2015;**70**(3):930-40. doi: 10.1093/jac/dku426 [published Online First: 2014/10/23]
20. Perrier M, Visseaux B, Landman R, *et al*. No impact of HIV-1 protease minority resistant variants on the virological response to a first-line PI-based regimen containing darunavir or atazanavir. *J Antimicrob Chemother* 2018;**73**(1):173-76. doi: 10.1093/jac/dkx366 [published Online First: 2017/10/28]
21. Mbisa JL, Kirwan P, Tostevin A, *et al*. Determining the origins of HIV-1 drug-resistant minority variants in people who are recently infected using phylogenetic reconstruction. *Clin Infect Dis* 2019;**69**(7):1136-43. doi: 10.1093/cid/ciy1048 [published Online First: 2018/12/12]
22. Casadella M, van Ham PM, Noguera-Julian M, *et al*. Primary resistance to integrase strand-transfer inhibitors in Europe. *J Antimicrob Chemother* 2015;**70**(10):2885-8. doi: 10.1093/jac/dkv202 [published Online First: 2015/07/19]
23. Doyle T, Dunn DT, Ceccherini-Silberstein F, *et al*. Integrase inhibitor (INI) genotypic resistance in treatment-naive and raltegravir-experienced patients infected with diverse HIV-1 clades. *J Antimicrob Chemother* 2015;**70**(11):3080-6. doi: 10.1093/jac/dkv243 [published Online First: 2015/08/28]
24. Stekler JD, McKernan J, Milne R, *et al*. Lack of resistance to integrase inhibitors among antiretroviral-naive subjects with primary HIV-1 infection, 2007-2013. *Antivir Ther* 2015;**20**(1):77-80. doi: 10.3851/IMP2780 [published Online First: 2014/05/17]
25. Alvarez M, Casas P, de Salazar A, *et al*. Surveillance of transmitted drug resistance to integrase inhibitors in Spain: implications for clinical practice. *J Antimicrob Chemother* 2019;**74**(6):1693-700. doi: 10.1093/jac/dkz067 [published Online First: 2019/03/07]
26. Karade S, Sen S, Sashindran VK. Absence of Integrase Strand Transfer Inhibitor Associated Resistance in Antiretroviral Therapy Naive and Experienced Individuals from Western India. *AIDS Res Hum Retroviruses* 2019;**35**(6):567-71. doi: 10.1089/AID.2018.0272 [published Online First: 2019/02/23]
27. Ji H, Patterson A, Taylor T, *et al*. Prevalence of Primary Drug Resistance Against HIV-1 Integrase Inhibitors in Canada. *J Acquir Immune Defic Syndr* 2018;**78**(1):e1-e3. doi: 10.1097/QAI.0000000000001649 [published Online First: 2018/02/10]
28. Charpentier C, Lee GQ, Rodriguez C, *et al*. Highly frequent HIV-1 minority resistant variants at baseline of the ANRS 139 TRIO trial had a limited impact on virological response. *J Antimicrob Chemother* 2015;**70**(7):2090-6. doi: 10.1093/jac/dkv048 [published Online First: 2015/03/11]

29. Margot NA, Kitrinos KM, Fordyce M, *et al*. Rare emergence of drug resistance in HIV-1 treatment-naive patients after 48 weeks of treatment with elvitegravir/cobicistat/emtricitabine/tenofovir alafenamide. *HIV Clin Trials* 2016;**17**(2):78-87. doi: 10.1080/15284336.2016.1142731 [published Online First: 2016/02/20]
30. Chang SY, Lin PH, Cheng CL, *et al*. Prevalence of Integrase Strand Transfer Inhibitors (INSTI) Resistance Mutations in Taiwan. *Sci Rep* 2016;**6**:35779. doi: 10.1038/srep35779 [published Online First: 2016/10/26]
31. Aghaizu A, Murphy G, Tosswill J, *et al*. Recent infection testing algorithm (RITA) applied to new HIV diagnoses in England, Wales and Northern Ireland, 2009 to 2011. *Euro Surveill* 2014;**19**(2) doi: 10.2807/1560-7917.es2014.19.2.20673 [published Online First: 2014/01/25]
32. Duong YT, Qiu M, De AK, *et al*. Detection of recent HIV-1 infection using a new limiting-antigen avidity assay: potential for HIV-1 incidence estimates and avidity maturation studies. *PLoS One* 2012;**7**(3):e33328. doi: 10.1371/journal.pone.0033328 [published Online First: 2012/04/06]
33. Kassanjee R, Pilcher CD, Busch MP, *et al*. Viral load criteria and threshold optimization to improve HIV incidence assay characteristics. *AIDS* 2016;**30**(15):2361-71. doi: 10.1097/QAD.0000000000001209 [published Online First: 2016/07/28]
34. Cunningham E, Chan YT, Aghaizu A, *et al*. Enhanced surveillance of HIV-1 drug resistance in recently infected MSM in the UK. *J Antimicrob Chemother* 2017;**72**(1):227-34. doi: 10.1093/jac/dkw404 [published Online First: 2016/10/16]
35. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**(15):2114-20. doi: 10.1093/bioinformatics/btu170 [published Online First: 2014/04/04]
36. Bankevich A, Nurk S, Antipov D, *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;**19**(5):455-77. doi: 10.1089/cmb.2012.0021 [published Online First: 2012/04/18]
37. Penedos AR, Myers R, Hadef B, *et al*. Assessment of the Utility of Whole Genome Sequencing of Measles Virus in the Characterisation of Outbreaks. *PLoS One* 2015;**10**(11):e0143081. doi: 10.1371/journal.pone.0143081 [published Online First: 2015/11/17]
38. Bennett DE, Camacho RJ, Otelea D, *et al*. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. *PLoS One* 2009;**4**(3):e4724. doi: 10.1371/journal.pone.0004724 [published Online First: 2009/03/07]
39. Tzou PL, Rhee SY, Descamps D, *et al*. Integrase strand transfer inhibitor (INSTI)-resistance mutations for the surveillance of transmitted HIV-1 drug resistance. *J Antimicrob Chemother* 2020;**75**(1):170-82. doi: 10.1093/jac/dkz417 [published Online First: 2019/10/17]
40. Gifford RJ, Liu TF, Rhee SY, *et al*. The calibrated population resistance tool: standardized genotypic estimation of transmitted HIV-1 drug resistance. *Bioinformatics* 2009;**25**(9):1197-8. doi: 10.1093/bioinformatics/btp134 [published Online First: 2009/03/24]
41. Abram ME, Ram RR, Margot NA, *et al*. Lack of impact of pre-existing T97A HIV-1 integrase mutation on integrase strand transfer inhibitor resistance and treatment outcome. *PLoS One* 2017;**12**(2):e0172206. doi: 10.1371/journal.pone.0172206 [published Online First: 2017/02/18]
42. Gupta S, Lataillade M, Kyriakides TC, *et al*. Low-frequency NNRTI-resistant HIV-1 variants and relationship to mutational load in antiretroviral-naive subjects. *Viruses* 2014;**6**(9):3428-37. doi: 10.3390/v6093428 [published Online First: 2014/09/27]
43. Nguyen T, Fofana DB, Le MP, *et al*. Prevalence and clinical impact of minority resistant variants in patients failing an integrase inhibitor-based regimen by ultra-deep sequencing. *J Antimicrob Chemother* 2018;**73**(9):2485-92. doi: 10.1093/jac/dky198 [published Online First: 2018/06/07]
44. World Health Organization. Update of recommendations on first- and second-line antiretroviral regimens. Geneva, Switzerland: 2019. Available at

https://apps.who.int/iris/bitstream/handle/10665/325892/WHO-CDS-HIV-19.15-eng.pdf?ua=1 (Accessed 04 November 2019).

45. Van Duyne R, Kuo LS, Pham P, *et al*. Mutations in the HIV-1 envelope glycoprotein can broadly rescue blocks at multiple steps in the virus replication cycle. *Proc Natl Acad Sci U S A* 2019;**116**(18):9040-49. doi: 10.1073/pnas.1820333116 [published Online First: 2019/04/13]

46. Malet I, Subra F, Charpentier C, *et al*. Mutations Located outside the Integrase Gene Can Confer Resistance to HIV-1 Integrase Strand Transfer Inhibitors. *MBio* 2017;**8**(5) doi: 10.1128/mBio.00922-17 [published Online First: 2017/09/28]

548 **Table 1.** Characteristics of the study population

Characteristic	*Category*	*Variable*[a]
Gender	Male	619 (94.5)
	Female	36 (5.5)
Risk exposure	MSM	543 (82.9)
	Heterosexual male	48 (7.3)
	Heterosexual female	33 (5.0)
	IDU	3 (0.5)
	Other / Unknown	28 (4.3)
Ethnicity	White	510 (77.9)
	Black (African / Caribbean / Other)	41 (6.3)
	Other / Unknown	104 (15.9)
Region	London	374 (57.1)
	North of England	113 (17.3)
	Midlands & East of England	83 (12.7)
	South of England	78 (11.9)
	Northern Ireland	6 (0.9)
	Wales	1 (0.2)
Subtype	A	24 (3.7)
	B	442 (67.5)
	C	38 (5.8)
	D	3 (0.5)
	F	27 (4.1)
	G	4 (0.6)
	CRF01_AE	28 (4.3)
	CFR02_AG	54 (8.2)
	CRF06_cpx	18 (2.7)
	CRF07_BC	1 (0.2)
	CRF11_cpx	3 (0.5)
	CRF12_BF	3 (0.5)
	CRF24_BG	1 (0.2)
	Complex recombinants	9 (1.4)
Age in years		33 [26.5-41]
CD4+ count (cells/mm^3)		545 [408-723][b]
Viral load (log_{10}, copies/mL)		5.18 [4.61-6.03][c]

549 [a]binary or category variables are presented as a number (% total) whereas continuous variables are
550 presented as median (interquartile range)
551 [b]n=602, number of samples with CD4+ counts done within 30 days of date of collection of the sample
552 used for RITA and sequencing
553 [c]n=366, number of samples with viral load done within 30 days of date of collection of the sample used for
554 RITA and sequencing
555

Figure Legends

Figure 1. *Flow chart showing study participant selection*. Patients were eligible for inclusion if they were newly diagnosed and identified to be recently infected (within 4 months of sampling) using RITA.

Figure 2. *The proportion of people on ART including InSTI in England between 2010 and 2015*. The data was estimated from ART prescribing data from NHS England and is stratified by PHE regions of London, Midlands & East of England, North of England and South of England. Data for London in 2015 does not include figures for the month of December. Hospitals in the former North West Strategic Health Authority (SHA) preferred to use RAL over bPIs as third agent for a significant portion of the survey period. Figures may include a source of over-estimation as RAL was used as first-line Post-Exposure Prophylaxis following Sexual Exposure (PEPSE). NA = data not available.

Figure 3. *Prevalence of InSTI RAMs in recently infected individuals in the UK between 2014 and 2016*. Stacked column graphs showing the prevalence of InSTI major and accessory resistance (A) and types of InSTI resistance mutations (B) at high (>20%) and low (2-20%) variant frequency. This was determined from 640 integrase sequences from recently infected individuals consisting of 316, 132 and 192 sequences in 2014, 2015 and 2016, respectively. Major and accessory InSTI resistance mutations were determined using the Stanford HIV Drug Resistance Database InSTI SDRM list 2019 and InSTI accessory mutations as of October 2019.

Figure 4. *Prevalence of PI and RTI TDR mutations in recently infected individuals in the UK between 2014 and 2016*. Stacked column graphs showing the prevalence of PI and RTI TDR mutations (A) and types of PI, NRTI and NNRTI resistance mutations (B) at high (>20%) and low (2-20%) variant frequency. This was determined from 622 protease and 597 RT sequences from recently infected individuals consisting of 302, 129 and 188 protease sequences and 295, 116 and 182 RT sequences in 2014, 2015 and 2016, respectively. PI, NRTI and NNRTI TDR mutations were determined using the WHO SDRM list 2019.

Figure 5. *Mutational load of low-frequency RAMs*. Box and whisker plot of mutational load of low frequency InSTI (n=14), PI (n=44), NRTI (n=28) and NNRTI (n=14) RAMs. Mutational load is defined as mutation frequency × viral load (copies/mL).

587

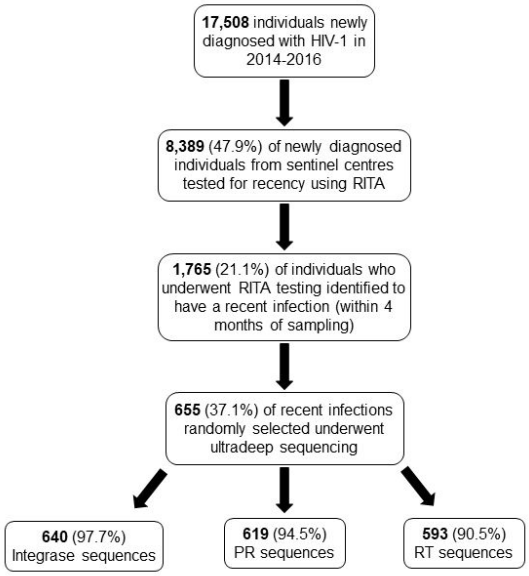

Figure 1

588

589

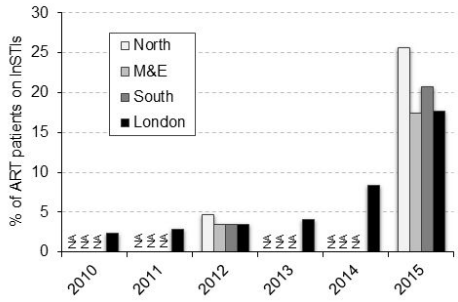

Figure 2

590

591

592

593

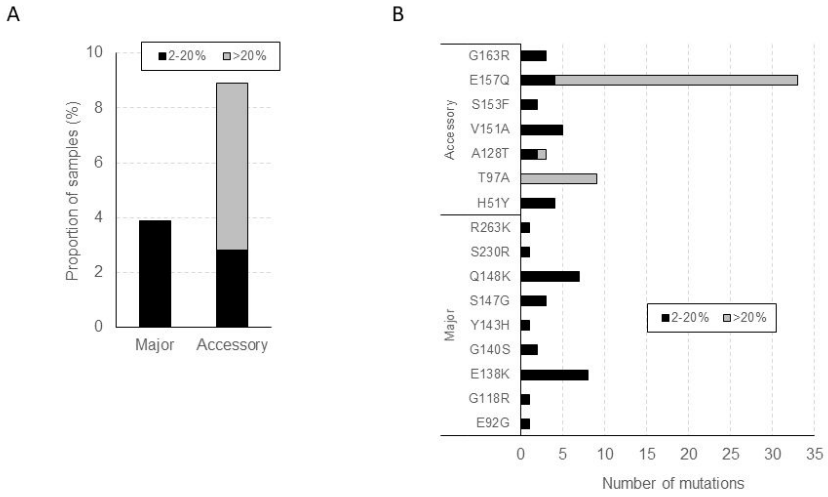

Figure 3

594

595

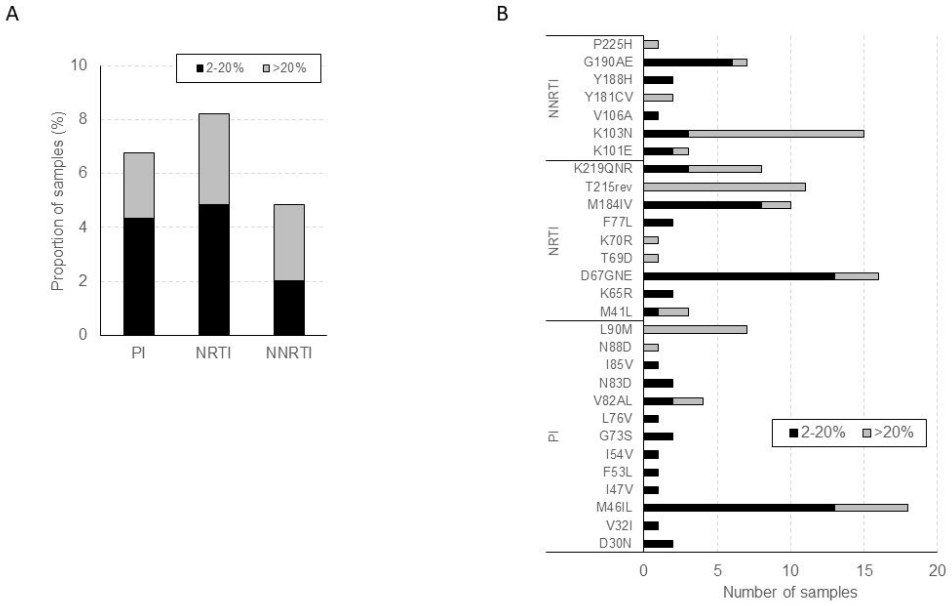

Figure 4

596

597

598

599

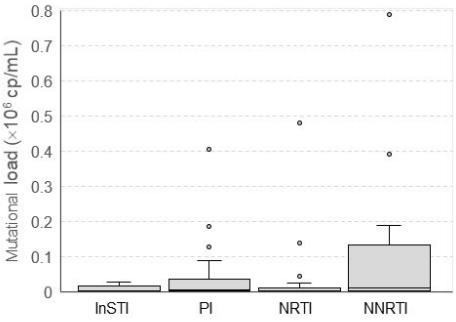

Figure 5

600

Table S1. Comparison of amplicon and sequence capture NGS-based methods

Variant frequency threshold	Total no. of bases[a]	Discordances			Positive percent agreement	95% CI
		Total	Absolute[b]	Partial[c]		
20%	15410	32	7	25	99.8	99.7-99.9
2%	15401	153	3	150	99.0	98.8-99.2

[a] using the integrase gene region from 19 different samples

[b] positions where the two methods gave completely different bases e.g. T vs C

[c] positions where one method had mixed bases whereas the other method had only one of the bases in the mixture e.g. Y vs T

Table S2. Analysis of mixed base calls between amplicon and sequence capture NGS-based methods

Variant frequency range	Method	No. of mixed base sites	No. concordant in alternate method	% concordant	No. concordant at >2% or 20%	% concordant overall
>20%	SeqCap	16	11	68.8	13	81.3
	Amplicon	29	10	34.5	21	72.4
2-20%	SeqCap	57	10	17.5	23	40.4
	Amplicon	116	11	9.5	14	12.1

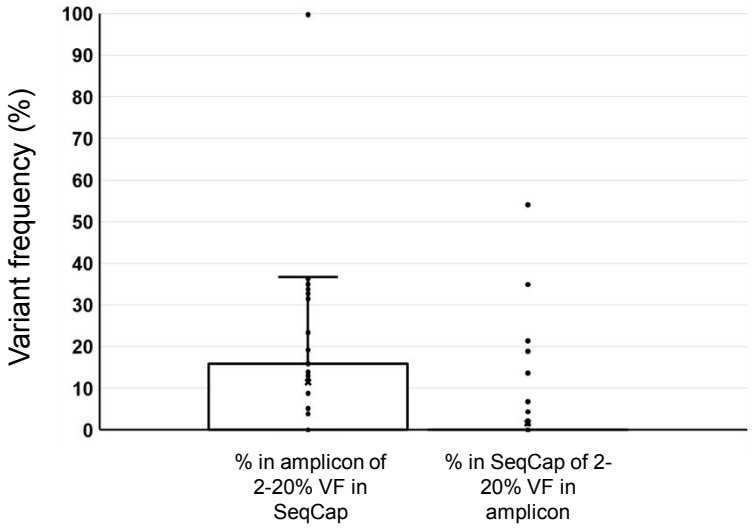

Figure S1. *Distribution of variants at mixed base positions*. The frequency of variants was determined for each method for variants detected within the 2-20% range by the alternate method and displayed in a box-whisker graph. This shows that low-frequency variants detected by the amplicon method are infrequently detected by the SeqCap method whereas the opposite is true.

Table S3. Codon frequency of low-frequency drug resistance variants

Gene	Sample ID	Mutation	AA:freq	Codon:freq	Depth
Integrase	14-0057	S147G	S:84.3487, G:15.4707,	AGT:84.0597, GGT:15.4587,	8306
Integrase	14-0128	E138K	E:92.2577, K:7.19498,	GAA:92.1551, AAA:7.18358,	8770
Integrase	14-0131	G140S	G:94.5347, S:5.05922,	GGC:94.3147, AGC:5.05922,	5910
Integrase	14-0144	Q148K	Q:97.4224, K:2.2942,	CAA:95.5735, AAA:2.2942, CAG:1.84885,	7410
Integrase	14-0159	Q148K	Q:97.3832, K:2.11281,	CAA:96.8599, AAA:2.11281,	5159
Integrase	14-0161	R263K	R:82.5018, K:16.3036,	AGA:82.2441, AAA:16.2567,	4269
Integrase	14-0186	Q148K	Q:97.2064, K:2.2843,	CAA:96.8394, AAA:2.2843,	13352
Integrase	14-0191	Q148K	Q:97.2017, K:2.54029,	CAA:96.9675, AAA:2.54029,	25194
Integrase	14-0261	Q148K	Q:97.1632, K:2.2705,	CAA:96.567, AAA:2.26505,	36732
Integrase	14-0280	Q148K	Q:95.7972, K:2.17376,	CAG:95.6766, AAG:2.17376,	29028
Integrase	14-0298	E138K	E:96.9016, K:2.32711,	GAA:96.6868, AAA:2.32711,	22818
Integrase	14-0302	S147G	S:87.8656, G:12.0877,	AGT:87.6945, GGT:12.0566,	6428
Integrase	14-0329	Q148K	Q:97.1253, K:2.34673,	CAA:96.7732, AAA:2.34673,	3409
Integrase	14-0395	S230R	S:96.964, R:2.06878,	AGC:96.8834, AGA:2.01505,	3722
Integrase	14-0416	E138K	E:89.5409, K:9.57476,	GAG:89.3389, AAG:9.5693,	18319
Integrase	14-0417	E92G	E:93.6004, G:5.76984,	GAG:93.4981, GGG:5.73835,	12704
Integrase	14-0419	E138K	E:95.5778, K:3.63601,	GAA:95.4206, AAA:3.62618,	10176
Integrase	14-0425	S147G	S:95.7666, G:4.15528,	AGT:95.5167, GGT:4.15528,	12803
Integrase	15-2792	E138K	E:96.3542, K:3.64583,	GAG:95.8333, AAG:3.64583,	384
Integrase	15-2828	E138K	E:91.4186, K:8.53367,	GAA:91.4026, AAA:8.53367,	6281
Integrase	15-3206	E138K	E:97.2211, K:2.50803,	GAA:97.2111, AAA:2.50803,	9968
Integrase	16-0742	G118R	G:96.7391, R:3.26087,	GGA:96.7391, AGA:3.26087,	92
Integrase	16-0775	Y143H	Y:97.2163, H:2.67666,	TAC:97.0021, CAC:2.67666,	934
Integrase	16-0830	G140S	G:93.5223, S:6.17409,	GGT:93.5223, AGT:6.17409,	1976
Integrase	16-0831	E138K	E:94.575, K:5.42504,	GAA:94.5265, AAA:5.42504,	4129
Integrase	14-0049	E157Q	E:94.3676, Q:5.38094,	GAA:94.3425, CAA:5.38094,	3977
Integrase	14-0103	V151A	V:89.7666, A:10.069,	GTA:89.2898, GCA:10.069,	12166
Integrase	14-0136	S153F	S:93.7058, F:5.66108,	TCT:93.5568, TTT:5.66108,	2685

Integrase	14-0174	G163R	G:96.2559, R:2.72596,	GGA:95.9735, AGA:2.72596,	15224
Integrase	14-0181	E157Q	E:94.8027, Q:4.96599,	GAA:94.5986, CAA:4.96599,	7350
Integrase	14-0181	G163R	G:97.969, R:2.00357,	GGA:97.5298, AGA:2.00357,	7287
Integrase	14-0379	G163R	G:95.9228, R:2.76765,	GGA:95.3827, AGA:2.76765,	7407
Integrase	14-0387	V151A	V:97.3938, A:2.4277,	GTG:96.769, GCG:2.4277,	5602
Integrase	14-0431	A128T	A:94.2695, T:5.68078,	GCC:94.1866, ACC:5.68078,	24134
Integrase	15-2838	H51Y	H:97.8102, Y:2.18978,	CAT:97.8102, TAT:2.18978,	137
Integrase	15-2861	V151A	V:93.573, A:3.15904, G:2.61438,	GTA:93.573, GCA:3.15904, GGA:2.61438,	918
Integrase	15-2942	V151A	V:95.4248, A:4.57516,	GTG:84.9673, GTA:10.4575, GCG:4.57516,	306
Integrase	15-2950	H51Y	H:92.2399, Y:7.58377,	CAT:92.2399, TAT:7.58377,	567
Integrase	15-2958	V151A	V:92.7253, A:4.10146,	GTA:92.6606, GCA:4.10146,	9265
Integrase	15-3064	S153F	S:85.4202, F:14.5798,	TCT:85.4202, TTT:14.5798,	583
Integrase	15-3252	E157Q	E:93.2653, Q:6.53061,	GAA:93.2653, CAA:6.53061,	980
Integrase	16-0739	H51Y	H:94.1402, Y:5.81341,	CAT:94.0938, TAT:5.81341,	8618
Integrase	16-0804	A128T	A:96.7692, T:3.20479,	GCA:96.7431, ACA:3.20479,	3838
Integrase	16-0804	V151A	V:93.5411, A:4.60812, M:1.75637,	GTG:93.4844, GCG:4.60812, ATG:1.75637,	5295
Integrase	16-0847	H51Y	H:97.764, Y:2.23602,	CAT:97.764, TAT:2.23602,	805
RT	14-0111	F77L	F:96.3229, L:2.47802, S:1.1191,	TTC:95.8433, CTC:2.47802, TCC:1.1191,	1251
RT	14-0128	D67N	D:94.9077, N:4.53875,	GAT:94.8155, AAT:4.5203,	5420
RT	14-0195	M41L	M:93.2409, L:5.60768,	ATG:93.2409, CTG:5.60768,	4690
RT	14-0218	K65R	K:96.8848, R:2.91979,	AAG:96.8197, AGG:2.91979,	9213
RT	14-0232	F77L	F:96.9987, L:2.43125,	TTC:96.9316, CTC:2.09591,	5964
RT	14-0282	K219R	K:94.1369, R:5.14612,	AAA:93.9826, AGA:5.14612,	11018
RT	14-0298	D67G	D:90.1886, G:9.1043,	GAC:90.0265, GGC:9.1043,	6788
RT	14-0311	D67G	D:96.3252, G:3.67483,	GAC:96.3252, GGC:3.67483,	898
RT	14-0353	D67G	D:97.2086, G:2.38127,	GAC:97.1689, GGC:2.38127,	7559
RT	14-0427	D67G	D:96.2117, G:3.06675,	GAC:96.0914, GGC:3.06675,	1663
RT	14-0434	K65R	K:97.5438, R:2.11851,	AAG:97.4823, AGG:2.11851,	3257
RT	14-0436	D67G	D:96.9207, G:2.69438,	GAC:96.6128, GGC:2.69438,	1299
RT	14-0437	M184I	M:95.5732, I:3.18471,	ATG:95.5732, ATA:2.92994,	3140

RT	14-0438	D67G	D:94.6474, G:5.09771,	GAC:94.4775, GGC:5.09771,	2354
RT	15-2840	D67N	D:97.3847, N:2.47003,	GAT:97.3847, AAT:2.47003,	2753
RT	15-2847	K219R	K:96.7532, R:3.24675,	AAA:96.7532, AGA:3.24675,	154
RT	15-3018	D67N	D:97.7273, N:2.27273,	GAT:97.7273, AAT:2.27273,	176
RT	15-3059	M184I	M:96.5116, I:3.48837,	ATG:96.5116, ATA:3.48837,	344
RT	15-3067	K219N	K:95.5437, N:4.45633,	AAA:95.5437, AAT:4.45633,	561
RT	15-3117	M184I	M:93.1483, I:6.8174,	ATG:93.1483, ATA:6.44056,	2919
RT	16-0741	M184I	M:96.5614, I:3.42205,	ATG:96.5614, ATA:3.33939,	6049
RT	16-0745	D67N	D:97.4684, N:2.53165,	GAT:97.4684, AAT:2.53165,	158
RT	16-0786	D67N	D:94.6404, N:5.35961,	GAT:94.6404, AAT:5.35961,	5075
RT	16-0799	M184I	M:91.1871, I:8.74101,	ATG:91.1871, ATA:8.70504,	2780
RT	16-0800	M184I	M:94.3534, I:5.64663,	ATG:94.3534, ATA:5.46448,	549
RT	16-0841	D67G	D:85.3253, G:14.5234,	GAC:85.3253, GGC:14.5234,	661
RT	16-0858	M184I	M:97.9458, I:2.05423,	ATG:97.9458, ATA:2.01315,	2434
RT	16-0879	M184I	M:96.0673, I:3.88411,	ATG:96.0673, ATA:3.83064,	20571
RT	16-0900	D67G	D:92.3077, G:7.50315,	GAT:92.1816, GGT:7.50315,	1586
RT	14-0080	V106A	V:97.2472, A:2.67103,	GTG:95.9389, GCG:2.61652,	3669
RT	14-0217	Y188H	Y:94.4268, H:5.25478,	TAT:94.2675, CAT:5.25478,	628
RT	14-0405	K103N	K:95.2284, N:3.98034,	AAA:95.1445, AAC:3.96835,	8341
RT	14-0443	K101E	K:96.0389, E:3.70736,	AAA:95.8416, GAA:3.70736,	7094
RT	14-0443	K103N	K:82.1165, N:16.9742,	AAA:81.9452, AAC:16.6974,	7588
RT	14-0443	Y188H	Y:92.4967, H:7.26386,	TAT:92.1951, CAT:7.25499,	11275
RT	15-2803	G190E	G:97.2282, E:2.69945,	GGA:97.2282, GAA:2.69945,	4149
RT	15-2837	K103N	K:89.7959, N:10.2041,	AAA:89.7959, AAC:10.2041,	441
RT	15-2861	G190E	G:95.8559, E:4.14414,	GGA:95.8559, GAA:4.14414,	1110
RT	15-3089	G190E	G:80, E:19.5918,	GGA:80, GAA:19.5918,	245
RT	15-3207	V106A	V:89.5053, A:10.4227,	GTA:89.5053, GCA:10.4227,	4164
RT	16-0745	G190E	G:92.9444, E:6.3772,	GGA:92.9444, GAA:6.3772,	737
RT	16-0769	G190E	G:96.5877, E:3.31754,	GGA:96.5877, GAA:3.31754,	2110
RT	16-0773	K101E	K:94.8252, E:5.17483,	AAA:94.8252, GAA:5.17483,	715

RT	16-0845	G190E	G:95.9751, E:3.85892,	GGA:95.9751, GAA:3.85892,	2410
PR	14-0077	M46I	M:97.0008, I:2.71159,	ATG:97.0008, ATA:1.97206,	2434
PR	14-0087	V82A	V:92.1878, A:7.6314,	GTC:91.2339, GCC:7.60646,	16039
PR	14-0128	M46I	M:94.0327, I:5.66519,	ATG:94.0327, ATA:5.47635,	10591
PR	14-0142	M46I	M:93.2759, I:6.37931,	ATG:93.2759, ATA:6.24521,	10440
PR	14-0196	I54V	I:97.2679, V:2.34663,	ATC:96.9913, GTC:2.34663,	11932
PR	14-0197	M46I	M:97.7584, I:2.18144,	ATG:97.7584, ATA:2.16639,	6647
PR	14-0215	M46I	M:92.0221, I:7.81991,	ATG:92.0221, ATA:7.74092,	2532
PR	14-0220	G73S	G:97.0144, S:2.83819,	GGT:96.6273, AGT:2.83819,	5426
PR	14-0224	N83D	N:95.1649, D:4.61019,	AAC:95.1274, GAC:4.61019,	5336
PR	14-0244	M46L	M:97.7225, L:2.05809,	ATG:97.7225, TTG:2.02674,	9572
PR	14-0295	N83D	N:96.67, D:3.14004,	AAC:96.6067, GAC:3.14004,	7898
PR	14-0323	I85V	I:94.8741, V:4.85126,	ATT:94.5538, GTT:4.85126,	2185
PR	14-0360	I47V	I:95.7078, V:3.93147,	ATA:95.6898, GTA:3.93147,	5545
PR	14-0399	M46I	M:96.6017, I:3.12115,	ATG:96.6017, ATA:3.01971,	55204
PR	14-0405	M46L	M:96.2874, L:3.18044,	ATG:96.2874, TTG:3.16776,	7892
PR	14-0419	M46I	M:85.6092, I:14.0484,	ATG:85.6092, ATA:13.5539,	7887
PR	14-0422	V32I	V:96.7893, I:3.00642,	GTA:96.6871, ATA:3.00642,	6852
PR	14-0444	D30N	D:88.5475, N:11.1034,	GAT:88.2542, AAT:11.1034,	7160
PR	15-2892	G73S	G:97.262, S:2.59003,	GGT:97.2373, AGT:2.59003,	4054
PR	15-3019	V82L	V:93.5012, L:6.39712,	GTC:93.4856, CTC:6.39712,	12787
PR	15-3050	F53L	F:84.6348, L:15.3652,	TTT:84.6348, CTT:15.3652,	397
PR	16-0717	M46I	M:91.4494, I:3.90283,	ATG:91.4494, ATA:3.88664,	6175
PR	16-0796	M46I	M:95.4325, I:4.56754,	ATG:95.4325, ATA:4.56754,	1029
PR	16-0901	L76V	L:96.1762, V:3.82381,	TTA:96.031, GTA:3.82381,	2066
PR	16-0992	M46I	M:86.7987, I:13.2013,	ATG:86.7987, ATA:13.2013,	303
PR	16-1001	D30N	D:95.2381, N:4.7619,	GAT:95.2381, AAT:4.7619,	147