MAJOR ARTICLE

Real-world prevalence of non-integrase INSTI resistance-associated mutations and virological outcomes in people who have recently acquired HIV-1 in the UK

*Christine Kelly[1,2], *James S. Lester[2,3], Daniel Bradshaw[1], David F. Bibby[1], Hodan Mohamed[1], Gary Murphy[1], Alison Brown[3], Caroline Sabin[4,5], Anna-Maria Geretti[6,7], Jean L. Mbisa[1,3,5]

[1]Virus Reference Department, UKHSA, Colindale, UK; [2]Centre for Experimental Pathogen Host Research, UCD, Ireland; [3]Blood Safety, Hepatitis, Sexually Transmitted Infections and HIV Division, UKHSA, London, UK; [4]Institute for Global Health, University College London, London, UK; [5]NIHR-HPRU in Bloodborne and Sexually Transmitted Infections at UCL, UK; [6]Department of Medicine of Systems, University of Rome Tor Vergata, Rome, Italy; [7]Department of HIV Medicine, Royal Free London NHS Trust (NorthMid), London, UK

Introduction: Integrase strand transfer inhibitors(INSTIs) are the mainstay of antiretroviral therapy(ART) globally. Virological breakthrough is uncommon but often manifests as low level viraemia and only 50% of cases have identified drug resistance mutations in the integrase gene. Non-integrase mutations in the Gag-nucleocapsid protein (NC), envelope glycoprotein (Env) and 3'polypurine tract (3'PPT) have been identified *in vitro*.

Methods: Between 2015 and 2021, HIV-1 whole genome sequencing was performed on samples from people with recently acquired HIV-1 in the UK. Sequences were linked to demographic and clinical data within the UK Health Security Agency's HIV and AIDS Reporting System. The relationship between non-integrase enzyme mutations and virological outcomes was assessed.

*Co-first authors

Correspondence: Christine Kelly: Catherine McAuley Centre, 41 Eccles Street, Dublin 7, Ireland christine.kelly@ucd.ie

375 (34%) of 1106 participants started an INSTI-based regimen. Of these, 337 (90%) were men and 196 (52%) were living with subtype B. The median age was 33 years and number of viral loads within 24 months of starting ART was 4.

Results: Overall, Env Y61H (33, 10%) and A539V (16, 5.0%), 3'PPT c9053t (17, 5.0%), and NC N8S (16, 4.8%) were the most prevalent non-integrase enzyme mutations. Univariable and multivariable Cox regression did not identify significant associations between the presence of these mutations individually and time to viral suppression, or to viral blip. Interestingly, accessory INSTI mutations were found significantly more frequently in people whose virus also harboured the Env mutation A539V (p=0.002).

Conclusion: Several non-integrase mutations were prevalent, but we found no evidence of an impact upon virological outcomes within treatment-naïve individuals on INSTI-based regimens who had recently acquired HIV.

Keywords: HIV; antiretroviral therapy; integrase inhibitors; INSTI; resistance; mutations

INTRODUCTION

Integrase strand transfer inhibitors (INSTIs) are now the mainstay of first line treatment for HIV-1 across the world[1]. In the UK, the British HIV Association guidelines recommend the use of the INSTIs dolutegravir and bictegravir as the additional agent in combination with nucleoside reverse transcriptase inhibitor (NRTI) treatment for initiation of therapy in adults living with HIV [2]. Part of the reason for their success is a high genetic barrier to resistance especially for the second-generation INSTIs that include dolutegravir, bictegravir and cabotegravir[3]. The latter is part of the only currently approved long-acting injectable regimen (cabotegravir/rilpivirine) which can also be used as a single agent for HIV preexposure prophylaxis (PrEP). However, INSTI resistance does arise and in studies of regimens containing dolutegravir in low and middle income countries, resistance in the integrase gene has been detected in 3.9-19.6% of individuals experiencing treatment failure [4]. In addition, INSTI resistance has been associated with non-integrase mutations in the 3'PPT [5], the envelope glycoprotein (Env) [6] and Gag-nucleocapsid protein (NC) [7], the clinical consequences of which are poorly defined. Within the viral life cycle, 3'PPT facilitates immune escape and infectivity of HIV-infected cells by modulating the expression of surface proteins [8], whilst Env is key to binding and fusion with host cells [9], and NC fulfils a number of roles, including facilitating viral assembly, chaperoning viral RNA, and stabilising viral DNA [10].

Mutations within the 3'PPT region have been implicated in dolutegravir resistance *in vitro*[5]. These consist of a replacement of the cytidine by a thymine in position 9053, and the modification of the subsequent GGGGGG sequence from position 9069 to 9073 to GCAGT, with a deletion in position 9073. Now unable to integrate, 1-LTR DNA circles are formed which

2

under the correct cellular conditions induce production of select viral proteins [11]. Although these viral products are unlikely to be a source of productive virus, they could contribute to detection of viral blips [12].

Several Env mutations have been implicated in dolutegravir resistance, including Y61H and P81S in gp120, and A539V and A556T in gp41[6]. Of these, virus with the A539V mutation in particular displays enhanced cell-to-cell transmission, unimpaired cell-free transmission, and significant reduction in sensitivity to dolutegravir and reduced sensitivity to other antiretroviral classes [13]. The other three mutations have been associated with the enhancement of cell-to-cell but impairment of cell-free transmission, alongside a reduction in sensitivity to dolutegravir. This enhancement of cell-to-cell transmission, seemingly mediated by the increased stability of Env conformations and decreased gp120 shedding, is thought to be the primary mechanism of dolutegravir resistance in these instances. *In vitro*, NC mutations, typically alongside Env mutations described above such as A539V, accelerate viral DNA integration leading to an exceptionally high multiplicity of infection (MOI) which can overwhelm the antiviral effect of INSTIs [7], [10].

Virological failure on a second generation INSTI-based treatment typically manifests as low-level viraemia and in the absence of drug resistance mutations within the integrase gene [3], [14]. This absence of recognised mutations raises questions as to the potential role for the described non-integrase mutations . Understanding if and how their detection should play a role in clinical practice is a critical element of future proofing this ART class.

Routinely, HIV drug resistance testing genotypes enzymes specific to anti-retroviral drug classes. Whole genome sequencing (WGS) has been provided as a reference surveillance service for a subset of new HIV diagnoses at the UKHSA national and WHO Global Specialised HIV Drug Resistance Laboratory in the UK since 2015. It affords the additional opportunity to examine resistance patterns outside of genotyped regions. Here, we link WGS data with clinical metadata to investigate the relationship between mutations that affect INSTI efficacy and virological outcomes.

METHODS

Population

UKHSA has sequenced HIV whole genomes from all individuals identified to have recently acquired HIV-1 in the UK since 2015 as part of routine HIV surveillance. All samples with sufficient leftover volume that have undergone avidity testing for recency of acquisition were eligible for WGS. Recent acquisition, defined as being in the preceding 6 months, is determined by the recent infection testing algorithm (RITA), which uses the results of an HIV-1 antibody avidity assay and clinical characteristics. The RITA algorithm excludes those with evidence of

treatment at time of sampling, and so the sequences to which WGS was applied were strictly pre-treatment. RITA results were available for approximately 21% of those newly diagnosed 2015-2021.

All individuals with available UKHSA WGS data, and linked clinical metadata from HARS, were eligible for inclusion. For the analysis of mutation prevalence, we excluded individuals whose RITA result did not confirm recent acquisition. For the analysis of time to viral suppression and blip we then further excluded those who did not start treatment on an INSTI-based regimen, had no viral load within 12 months of starting treatment, or had a missing or suppressed viral load at baseline. For the analysis of time to viral blip, we further excluded those who were no longer on an INSTI-based regimen at time of viral suppression, those who did not have viral loads available after viral suppression, and those who did not reach viral suppression within 24 months of starting treatment.

Laboratory methods

The presence of clinically relevant drug resistance mutations (surveillance drug resistance mutations, SDRMs) was identified by submitting sequences to the Stanford Calibrated Population Resistance (CPR) tool[15]. This tool both identifies established SDRMs and applies sequence and mutation-level exclusion criteria to remove potentially spurious findings based on factors including low coverage, multiple APOBEC mutations, or adjacency to insertions, deletions or frame shifts [16], [17]. We also identified accessory INSTI resistance mutations using the Stanford HIVdb program to assess the co-occurrence of known integrase mutations with the mutations of interest [18]. Subtypes were obtained using the Stanford HIVdb program and the COMET HIV-1 subtyping tool[18], [19]. Concordance between the two methodologies was high (99.6%), with Stanford able to assign a higher portion of samples to a subtype (93.7% vs 66.3%), and thus was used for subsequent analysis.

The complete genomic sequencing of HIV was performed using a previously described sequence capture method and the short read data was assembled into consensus whole genomes using Genomancer, an in-house viral genome assembly pipeline [20].

Consensus sequences derived at 30x depth and 20% variant frequency threshold were aligned against the HXB2 reference sequence using MAFFT, and mutations of interest were then identified either within the aligned 3'PPT nucleotide sequence or translated Env or NC amino acid sequence as appropriate[21]. Where reads at a given locus were ambiguous, mutation presence was considered unknown. Throughout this manuscript we have used position numbering consistent with how these mutations were originally described, rather than their actual HXB2 positions, but it should be noted that Env 209, 539 and 556 are instead 211, 541 and 558 in HXB2, and thus our alignments. Similarly, the 3'PPT positions of interest referred to as 9053, 9069, 9070, 9072 and 9073 correspond to 9063, 9079, 9080, 9082 and 9083. NC positions were unchanged.

4

Outcomes

Our primary outcome was a viral blip defined as a viral load greater than 50 after having reached viral suppression. Although the aetiology of detectable viral loads on ART is complex and multi-factorial, the occurrence of viral blips carries an increased risk of future virological failure [22]. Because virological failure is now a rare occurrence, we selected viral blip as a pragmatic virological outcome. For this outcome people were followed from viral suppression whilst on an INSTI-based treatment, to date of death, change to a non-INSTI-based regimen, or final available viral load. A viral load threshold of 50 was used for these outcome measures to account for INSTI resistance often manifesting as persistent low-level viraemia. Data across all available time periods for included people were used.

The secondary outcome was time to viral suppression post ART initiation defined as time from the start of treatment to the first VL ≤50 copies/mL. For this outcome, people were followed from ART initiation with an INSTI-based treatment to date of death, change to a non-INSTI-based regimen (including both change to a PI-based regimen, or the addition of a PI to the regimen), or final available viral load. Follow-up was limited to 24 months after ART initiation.

Variable management

Antiretroviral therapy (ART) start date, baseline VL, subsequent VL, and ART regimen all required derivation from HARS.

ART start date

ART initiation is captured by both explicit ART start dates, and ART status at attendances. For the purposes of analysis, ART start date taken to be the earliest of either a reported ART start date, or the first of two consecutive attendances where an individual was reported to be on treatment. If an individual had evidence of viral suppression before this date, treatment initiation was backdated to date of baseline VL collection.

Baseline VL

Where a VL was available at ART initiation, this was used. If not available, the highest VL within 30 days before or 7 days after was used, otherwise the nearest value within 1 year before starting ART, otherwise the nearest within 1 week after, or finally the nearest value at any time before starting ART.

Viral load

Viral load was occasionally ambiguously reported as 0 to indicate either an undetectable viral load, or where viral load was not tested, and as such was treated as missing data. Where a VL was reported as being between 1 and 5, this was assumed to have been log transformed, and was corrected accordingly. Repeated identical VL values > 100 copies/ml were assumed to reflect the

results of previous tests which had been carried forwards, with subsequent duplicated values treated as missing data.

ART regimen

When available, an individual's starting regimen was taken to be their first recorded ART regimen within 1 year of ART initiation, provided it contained no ambiguous medications (e.g. other unknown, other protease inhibitor). When ambiguity was present within this first recorded regimen, clarifications within 1 year of starting ART were taken to reflect this first regimen. Where multiple distinct regimens were listed on the same day, the more complete regimen was used. If both were equally complete, they were merged.

Statistical analysis

To identify associations between the mutations of interest and individual and virological characteristics, we used the Wilcoxon test for continuous variables, the Chi-squared test for categorical variables where all expected values were ≥ 5, and Fisher's exact test otherwise. We used univariable and multivariable Cox regression to identify any associations between time to viral suppression and viral blip and the mutations of interest. Univariable Cox regression was used to identify variables for inclusion in a multivariable regression, with all variables with a $p < 0.1$, and key characteristics (age and gender) included in the multivariable regression.

Ethics

Ethical approval was obtained through the UKHSA Research Ethics and Governance Group (REGG) on 05/04/2022 (reference number NR0303) and the UKHSA Caldicott Advisory Panel on 29/11/2022 (reference number CAP-2022-18).

RESULTS

Population identification

1182 people had available WGS data between 2015-2021. Of these, 76 HIV acquisitions could not be confirmed to be recent (figure 1).

Sequentially excluding individuals who did not start on an INSTI-based regimen, or whose first regimen was unknown (n=731), those with no VL within 12 months of ART initiation (n=29), and those who were virally suppressed at baseline (n=79) left 267 people suitable for analysis of time to viral suppression, our secondary outcome. A further 26 people were excluded from the analysis of time to subsequent viral blip, our primary outcome, due to switching to a non-INSTI-based regimen before reaching viral suppression (n=10), having no recorded VL after viral suppression (n=13), or no viral suppression (n=3) within 24 months of starting treatment. Of

those with no viral suppression within 24 months, none had evidence of any of the mutations of interest.

Demographic and clinical characteristics

The overall population comprised 990 (90%) men, and 806 (75%) people of white ethnicity (table 1). The most common probable acquisition route (816; 78%) was sex between men. The majority had been born in the UK (655; 63%) and were living with HIV subtype B (610; 55%). No important demographic differences were noted between the overall sample and those who were on an INSTI-based regimen. Most people had a baseline CD4 above 350 cells/μL (862; 79%) and an initial HIV VL of ≥100,000 copies/mL (483; 55%).

Anti-retroviral therapy

Those who started on an INSTI-based regimen had typically started their treatment more recently than those starting other regimens, and had done so on a raltegravir (205, 55%) or dolutegravir (137, 37%) based regimen, with smaller portions of the sample receiving elvitegravir (5, 1.3%) or bictegravir (28, 7.5%). Thus, 56% started on a first generation INSTI, and 44% on a second generation INSTI. Overall, individuals had a median of 5 viral loads available within 24 months of ART initiation, and a median duration of follow-up of 19 months. This was consistent across all treatment groups. Further details on major drug resistant mutations can be found in supplementary tables.

Genotypic resistance prevalence

NRTI SDRMs were most prevalent (44; 4.1%), followed by NNRTI SDRMs (35; 3.2%), and PI SDRMs (29; 2.7%). INSTI SDRMs were seen in less than 1% of sequences (detailed breakdown available in supplementary material).

Non-integrase enzyme INSTI resistance mutations

Of the 3'PPT, Env and nucleocapsid mutations of interest, 3'PPT c9053t (47; 4.7%), Env Y61H (102; 11%), Env A539V (36; 3.9%), and nucleocapsid N8S (52, 5.2%) were most prevalent.

The presence of 3'PPT 9053t was significantly associated with both region of birth and viral subtype (see table 2). Env Y61H was significantly associated with a lower baseline viral load, and absence of INSTI accessory mutations. Env A539V was significantly associated with age at ART initiation, ethnic group, region of birth, baseline viral load and viral subtype. Individuals carrying a virus with this mutation were typically older (median age 38 vs 32), more commonly of an ethnicity other than White (43% vs 25%),born outside the UK (59% vs 36%), were more likely to have a non-B subtype (86% vs 44%), and more often had at least one INSTI accessory mutation (22% vs 7.8%). NC N8S was significantly associated with younger age (median age 30 vs 32) and viral subtype B. Two of the eight participants with a HIV virus displaying INSTI SDRMs also had this mutation.

7

Multivariable analysis of associations between non-integrase enzyme INSTI mutations and virological outcomes

For the analysis of time to viral suppression there were 103 person-years of follow-up, and the median time to detecting viral suppression was 89 days. Of 267 individuals who started treatment, 248 had reached viral suppression within 24 months, 8 switched regimens before reaching viral suppression, and 11 did not have evidence of viral suppression before their final VL. For viral blip, a total of 66 events were observed over 815 person-years of follow-up. Univariable Cox analysis identified female gender, lower baseline CD4, and higher baseline VL as associated with a longer time to viral suppression, and female as associated with a shorter time to viral blip. No significant effect was observed for any of the non-integrase mutations under consideration in univariable analysis for time to viral suppression [hazard ratio (HR), 95% confidence interval (CI): 3'PPT c9053t: 1.31, (0.74, 2.31); Env Y61H: 1.18, (0.79, 1.74); Env A539V: 0.88, (0.49, 1.58); NC N8S: 0.73, (0.42, 1.26)], or for time to viral blip [HR, (CI): 3'PPT c9053t: 1.44, (0.52, 3.97); Env Y61H: 0.66, (0.28, 1.54); Env A539V: 0.83, (0.26, 2.67); NC N8S: 0.57, (0.18, 1.82)]. These results remained non-significant after adjustment for other variables (Table 3). Two sub-analyses were performed, one limiting analysis of both outcomes only to those on second generation INSTIs, yielding consistent results, and a second censoring those with possible virological failure after viral blip (VL>50 following blip), again yielding consistent results (supplementary materials).

DISCUSSION

We have identified 3'PPT c9053t, Env Y61H,Env A539V, and NC N8S at approximately 5% and higher prevalence within this sample. Three of these mutations displayed significant associations with subtype, with two also significantly associated with region of birth in univariable analysis. This is unsurprising given people born abroad who are diagnosed with HIV in the UK have often acquired their HIV prior to arrival, and thus the subtype of their HIV reflects local circulating subtypes. Marked difference in the prevalence of these mutations by subtype, which may in part reflects natural polymorphisms, are exemplified by A539V where a majority of sequences with this mutation were subtype CRF02_AG. However, analysis of the 'Web alignments' from the Los Alamos HIV database using the AnalyzeAlign tool indicated that the A539V mutation is seen in the majority of subtype G sequences, whereas the same is not seen of CRF02_AG itself (supplementary materials), in contrast to what we observed. This may reflect an aspect of diversity within this subtype which is not currently captured by the sequences included, or overrepresented within our sample.

We identified a low prevalence of transmitted INSTI SDRMs, with a relevant mutation seen in <1% of all suitable sequences, but a higher prevalence of integrase accessory mutations. The presence of A539V was significantly associated with INSTI accessory mutations, which is of

interest given the putative role of A539V as a facilitator for the accumulation of INSTI resistance mutations.

We did not see any significant impact of the non-integrase mutations considered, either in univariable nor multivariable analysis, on time to virological suppression or viral blip. Though these mutations have been shown to play a role directly, or in facilitating INSTI resistance in vitro, this was not replicated in vivo within our sample. Hikichi et al have recently demonstrated through in vitro experiments that A539V was one of the first Env mutations to appear during serial passage under increasing concentrations of dolutegravir, conferring 6 fold resistance to dolutegravir [7]. Accumulation of further Env mutations was necessary to cause high level resistance with >2000-fold resistance demonstrated in viruses containing seven mutations. Therefore, a series of Env mutations may be necessary for clinically detectable viral load. However, A539V viruses do exhibit a replication advantage compared to wild type and may be an indication of risk of accumulation of further Env mutations in individuals with suboptimal adherence over longer periods of time. This is somewhat substantiated by the co-existence of A539V with known INSTI accessory mutations here, although the numbers were too small to examine any virological impact. Follow-up studies are required to assess the dynamics of Env mutation accumulation for patients with A539V viruses at treatment initiation. Investigation is also required on the extent to which cell-to-cell transfer contributes to fitness advantage in Env mutation viruses, to what extent this process can contribute to a detectable viral load, and whether it can drive accumulation of further Env mutations.

HARS is a data set for national HIV surveillance, and so whilst it provides nationally comprehensive and detailed longitudinal data, certain details are not requested, or subject to the real-world context of HIV care. As a result, significant heterogeneity exists in characteristics such as exactly when viral loads are collected, and valuable predictors such as adherence are not routinely collected. Cleaning of viral loads included correcting potentially log transformed values. We note that our conclusions were unchanged when we experimentally excluding these values entirely. Missingness was present in recorded antiviral regimens, which was handled conservatively. The number of individuals with the mutations of interest was relatively small in this pre-treatment population, precluding our ability to disentangle effects of mutation from subtype. Poor virological outcomes were uncommon, reflecting both the efficacy of INSTI-based ART, and high-quality HIV care received, but potentially impairing the statistical power of our analysis. Further, although viral blips are predictive of subsequent virological failure, a minority go on to experience this outcome[22]. Viral blips are multi-factorial in aetiology and may be driven by factors other than resistance, such as large viral reservoir, infection driven homeostatic proliferation, and isolated adherence issues[22], [23]. Therefore, viral blips for patients studied here do not necessarily imply drug resistance. However, all these factors may increase the risk of development of future resistance, and virological failure, potentiated by the presence of non-integrase INSTI resistance mutations in pre-treatment viruses.

These data justify further work investigating the role of non-integrase mutations in the development of INSTI resistance driving virological and clinical failure. Future studies could investigate pre-treatment and on-treatment WGS longitudinally, and ideally over a longer time period. This could be coupled with in vitro phenotypic susceptibility testing using replication competent and subtype-specific recombinant virus systems. Of particular interest may be the mechanisms through which various non-integrase mutations confer resistance in vivo, and the extent to which viral blips are indicative of the presence of replication competent virus.

Our results indicate that the presence of individual mutations outside the integrase gene do not have a significant impact on viral blips or time to suppression when present at ART initiation. Relationships between integrase gene and non-integrase gene mutations are interesting and warrant further study.

Author contributions: Analysis conceived by DB, AMG, CK, JSL, JLM and CS, data collected and contributed by AB, GM, and HM, analysis performed by DFB, CK and JSL, paper written by CK and JSL.

References

[1]   Consolidated Guidelines on HIV Prevention, Testing, Treatment, Service Delivery and Monitoring: Recommendations for a Public Health Approach, 1st ed. Geneva: World Health Organization, 2021.

[2]   L. Waters et al., 'BHIVA guidelines on antiretroviral treatment for adults living with HIV-1 2022', HIV Med., vol. 23, no. S5, pp. 3–115, Dec. 2022, doi: 10.1111/hiv.13446.

[3]   A. V. Zhao, R. D. Crutchley, R. C. Guduru, K. Ton, T. Lam, and A. C. Min, 'A clinical review of HIV integrase strand transfer inhibitors (INSTIs) for the prevention and treatment of HIV-1 infection', Retrovirology, vol. 19, no. 1, Art. no. 1, Dec. 2022, doi: 10.1186/s12977-022-00608-1.

[4]   'HIV drug resistance – brief report 2024', WHO. Accessed: Mar. 08, 2024. [Online]. Available: https://iris.who.int/bitstream/handle/10665/376039/9789240086319-eng.pdf?sequence=1

[5]   I. Malet et al., 'Mutations Located outside the Integrase Gene Can Confer Resistance to HIV-1 Integrase Strand Transfer Inhibitors', mBio, vol. 8, no. 5, pp. e00922-17, Sep. 2017, doi: 10.1128/mBio.00922-17.

[6]   R. Van Duyne, L. S. Kuo, P. Pham, K. Fujii, and E. O. Freed, '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., vol. 116, no. 18, pp. 9040–9049, Apr. 2019, doi: 10.1073/pnas.1820333116.

[7]   Y. Hikichi, J. R. Grover, A. Schäfer, W. Mothes, and E. O. Freed, 'Epistatic pathways can drive HIV-1 escape from integrase strand transfer inhibitors', Sci. Adv., vol. 10, no. 9, p. eadn0042, Mar. 2024, doi: 10.1126/sciadv.adn0042.

[8]   C. Z. Buffalo, Y. Iwamoto, J. H. Hurley, and X. Ren, 'How HIV Nef Proteins Hijack Membrane Traffic To Promote Infection', J. Virol., vol. 93, no. 24, pp. e01322-19, Dec. 2019, doi: 10.1128/JVI.01322-19.

[9]   K. T. Arrildt, S. B. Joseph, and R. Swanstrom, 'The HIV-1 env protein: a coat of many colors', Curr. HIV/AIDS Rep., vol. 9, no. 1, pp. 52–63, Mar. 2012, doi: 10.1007/s11904-011-0107-3.

[10]   Y. Hikichi et al., 'Elucidating the Mechanism by Which HIV-1 Nucleocapsid Mutations Confer Resistance to Integrase Strand Transfer Inhibitors', May 18, 2025, bioRxiv. doi: 10.1101/2025.05.17.654662.

[11]   C. Richetta et al., 'Mutations in the 3'-PPT Lead to HIV-1 Replication without Integration', J. Virol., vol. 96, no. 14, pp. e00676-22, Jul. 2022, doi: 10.1128/jvi.00676-22.

[12]   J. G. Dekker, B. Klaver, B. Berkhout, and A. T. Das, 'HIV-1 3′-Polypurine Tract Mutations Confer Dolutegravir Resistance by Switching to an Integration-Independent Replication Mechanism via 1-LTR Circles', J. Virol., pp. e00361-23, May 2023, doi: 10.1128/jvi.00361-23.

[13]   Y. Hikichi et al., 'Mechanistic Analysis of the Broad Antiretroviral Resistance Conferred by HIV-1 Envelope Glycoprotein Mutations', mBio, vol. 12, no. 1, pp. e03134-20, Feb. 2021, doi: 10.1128/mBio.03134-20.

[14]   C. Chu et al., 'Prevalence of Emergent Dolutegravir Resistance Mutations in People Living with HIV: A Rapid Scoping Review', Viruses, vol. 16, no. 3, p. 399, Mar. 2024, doi: 10.3390/v16030399.

[15]   R. J. Gifford et al., 'The calibrated population resistance tool: standardized genotypic estimation of transmitted HIV-1 drug resistance', Bioinformatics, vol. 25, no. 9, pp. 1197–1198, May 2009, doi: 10.1093/bioinformatics/btp134.

[16] D. E. Bennett et al., 'Drug Resistance Mutations for Surveillance of Transmitted HIV-1 Drug-Resistance: 2009 Update', PLoS ONE, vol. 4, no. 3, p. e4724, Mar. 2009, doi: 10.1371/journal.pone.0004724.

[17] P. L. Tzou et al., 'Integrase strand transfer inhibitor (INSTI)-resistance mutations for the surveillance of transmitted HIV-1 drug resistance', J. Antimicrob. Chemother., vol. 75, no. 1, pp. 170–182, Jan. 2020, doi: 10.1093/jac/dkz417.

[18] T. F. Liu and R. W. Shafer, 'Web Resources for HIV Type 1 Genotypic-Resistance Test Interpretation', Clin. Infect. Dis., vol. 42, no. 11, pp. 1608–1618, Jun. 2006, doi: 10.1086/503914.

[19] D. Struck, G. Lawyer, A.-M. Ternes, J.-C. Schmit, and D. P. Bercoff, 'COMET: adaptive context-based modeling for ultrafast HIV-1 subtype identification', Nucleic Acids Res., vol. 42, no. 18, pp. e144–e144, Oct. 2014, doi: 10.1093/nar/gku739.

[20] J. L. Mbisa et al., 'Surveillance of HIV-1 transmitted integrase strand transfer inhibitor resistance in the UK', J. Antimicrob. Chemother., vol. 75, no. 11, pp. 3311–3318, Nov. 2020, doi: 10.1093/jac/dkaa309.

[21] K. Katoh and D. M. Standley, 'MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability', Mol. Biol. Evol., vol. 30, no. 4, pp. 772–780, Apr. 2013, doi: 10.1093/molbev/mst010.

[22] O. Elvstam et al., 'Virologic Failure Following Low-level Viremia and Viral Blips During Antiretroviral Therapy: Results From a European Multicenter Cohort', Clin. Infect. Dis., vol. 76, no. 1, pp. 25–31, Jan. 2023, doi: 10.1093/cid/ciac762.

[23] N. Bachmann et al., 'Determinants of HIV-1 reservoir size and long-term dynamics during suppressive ART', Nat. Commun., vol. 10, no. 1, p. 3193, Jul. 2019, doi: 10.1038/s41467-019-10884-9.

Figure 1. caption: Inclusion criteria flow chart

Alt text: Graphical representation of people who were included within the analysis after the successive exclusion of subsets of people according to inclusion criteria

Figure 1. Inclusion flow chart

People who initiated ART 2015-2021 with a whole genome sequence
n = 1182

Not confirmed to be recent
n = 76

Suitable for analysis of mutation prevalence
n = 1106

Not started on INSTI-based regimen, or first regimen unknown
n = 731

People whose first regimen was INSTI-based
n = 375

No VL within 12 months of ART initiation
n = 29

Virally suppressed at baseline
n = 79

Suitable for analysis of time to viral suppression
n = 267

• Changed to a non-INSTI regimen before viral suppression
• No VL after viral suppression
• Not virally suppressed within 24 months of ART initiation
n = 26

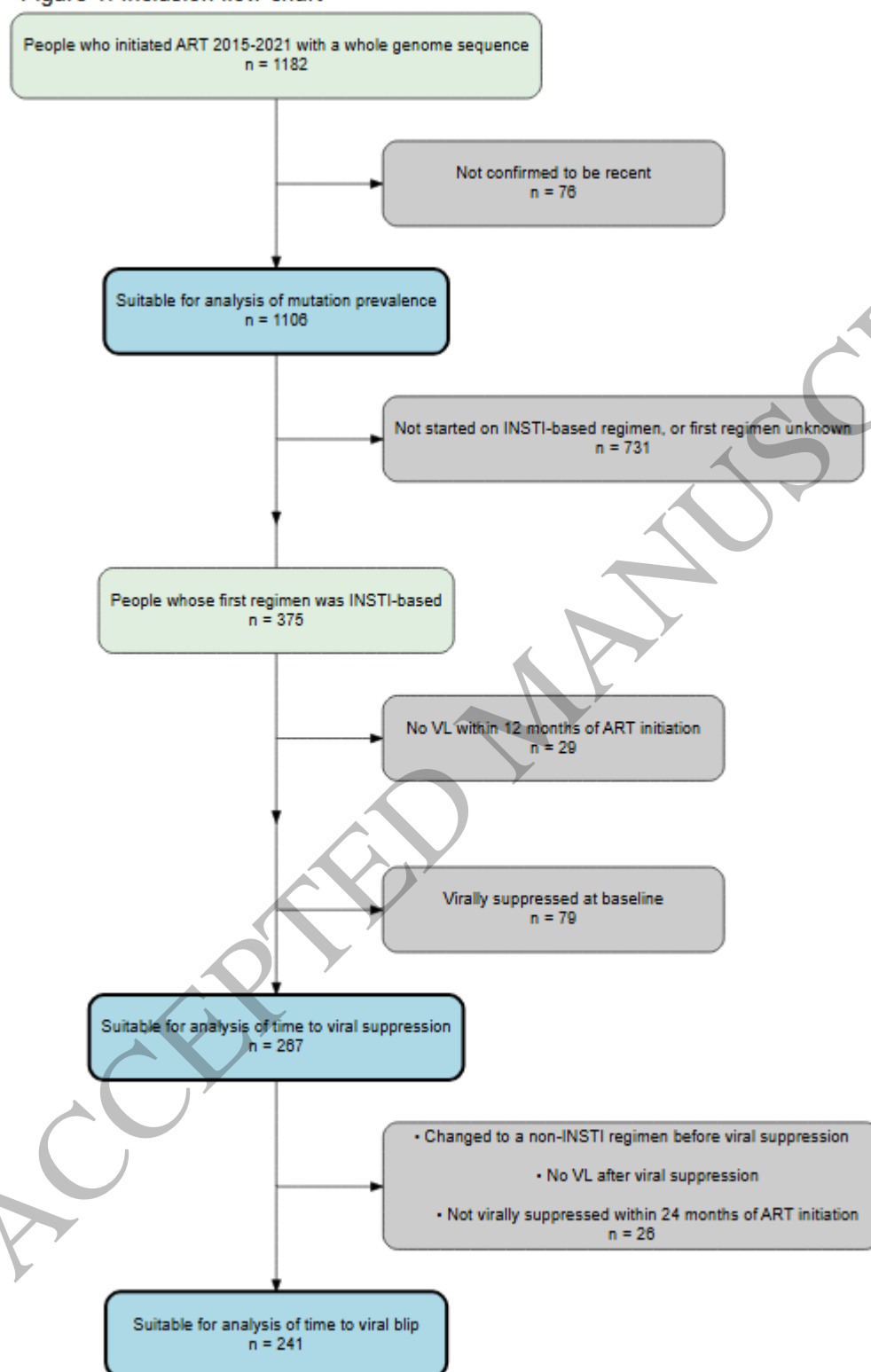Suitable for analysis of time to viral blip
n = 241

Figure 2. Caption: Kaplan-Meier curves of time to viral suppression (A) and viral blip (B), stratified by the presence of 3'PPT c9053t.

Alt text: Graphs of Kaplan-Meier curves of time to viral suppression and time to viral blip, stratified by the presence of 3'PPT c9053t with subfigures labelled from A to B, illustrating associations between these mutations and outcomes.

Figure 2. Kaplan-Meier curves for clinical outcomes by 3'PPT mutation presence

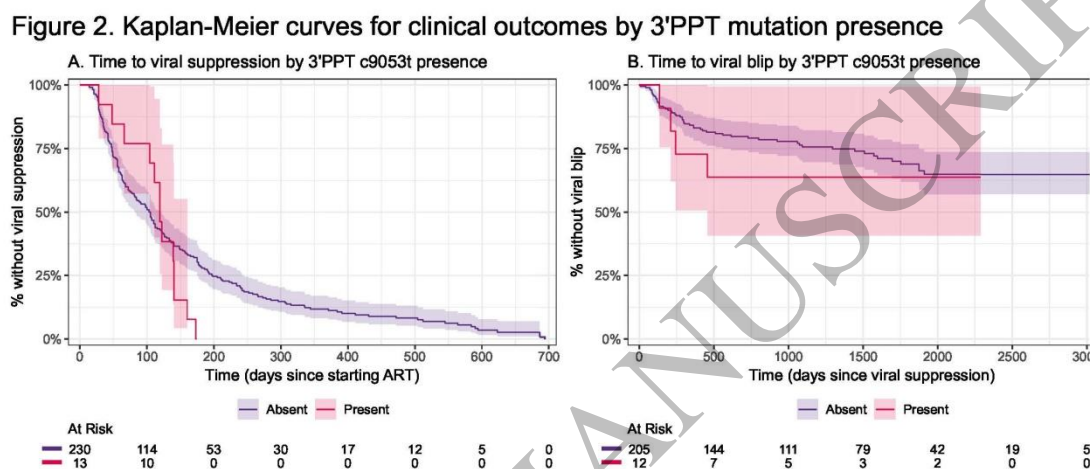**Figure 3.** Caption: Kaplan-Meier curves of time to viral suppression (A,C) and viral blip (B,D), stratified by the presence of Env Y61H (A,B) and Env A539V (C,D).

Alt text: Graphs of Kaplan-Meier curves of time to viral suppression and time to viral blip, stratified by the presence of Env Y61H and Env A539V with subfigures labelled from A to B, illustrating associations between these mutations and outcomes.

Figure 3. Kaplan-Meier curves for clinical outcomes by Env mutation presence

Figure 4. Caption: Kaplan-Meier curves of time to viral suppression (A) and viral blip (B), stratified by the presence of NC N8S.

Alt text: Graphs of Kaplan-Meier curves of time to viral suppression and time to viral blip, stratified by the presence of NC N8S with subfigures labelled from A to B, illustrating associations between these mutations and outcomes.

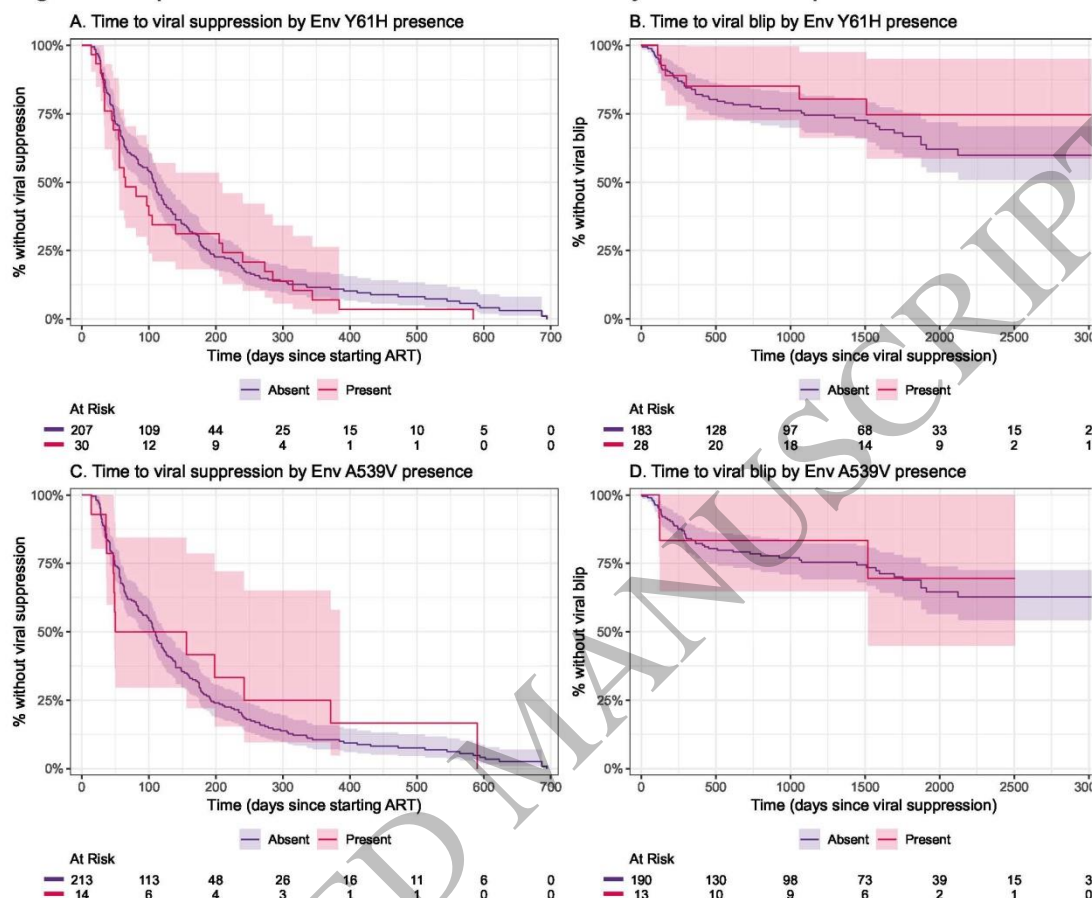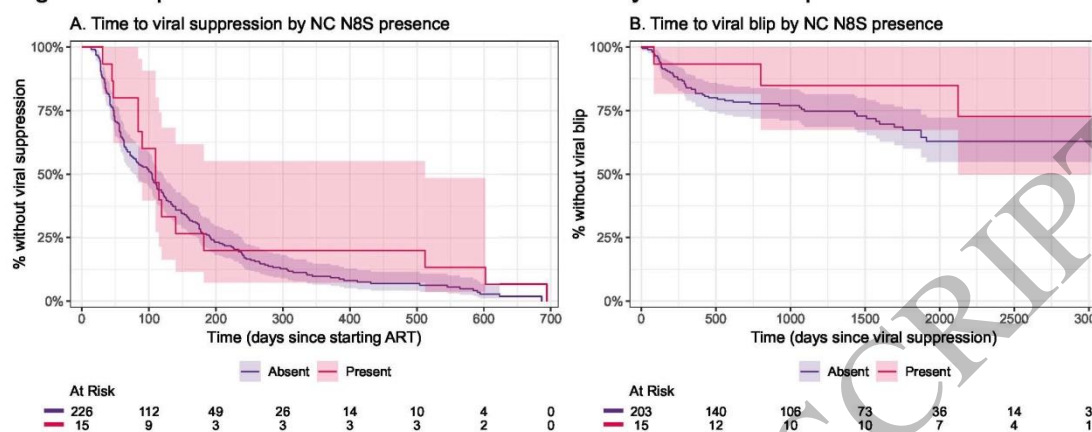Figure 4. Kaplan-Meier curves for clinical outcomes by NC mutation presence

Table 1: Baseline characteristics of the study population

Characteristic	Started on an INSTI-based regimen N = 375	Started on any regimen (including unknown) N = 1,106

Characteristic	Started on an INSTI-based regimen N = 375	Started on any regimen (including unknown) N = 1,106
Age at ART start, Median (Q1, Q3)	33 (26, 41)	32 (26, 42)
Gender, n (%)		
Men (including trans men)	337 (90%)	990 (90%)
Women (including trans women)	38 (10%)	115 (10%)
Non-Binary	0 (0%)	1 (<0.1%)
Ethnic group, n (%)		
White	269 (74%)	806 (75%)
Black	35 (9.6%)	101 (9.4%)
Asian	28 (7.7%)	71 (6.6%)
Other/Mixed	32 (8.8%)	95 (8.9%)
Unknown	11	33
Probable route of HIV acquisition, n (%)		
Sex between men	265 (77%)	816 (78%)
Sex between men and women	75 (22%)	220 (21%)
Other	3 (0.9%)	10 (1.0%)
Unknown	32	60
Region of birth, n (%)		
United Kingdom	210 (60%)	655 (63%)
Africa	27 (7.7%)	74 (7.1%)
Americas	20 (5.7%)	60 (5.7%)
Asia & Oceania	33 (9.4%)	74 (7.1%)
Europe	62 (18%)	183 (17%)
Unknown	23	60
Baseline CD4 (cells/µl), n (%)		
<350 cells/µl	73 (20%)	227 (21%)
≥350 cells/µl	290 (80%)	862 (79%)
Unknown	12	17
Baseline VL (copies/mL), n (%)		
<100,000 copies/mL	132 (45%)	394 (45%)
≥100,000 copies/mL	160 (55%)	483 (55%)

Characteristic	Started on an INSTI-based regimen N = 375	Started on any regimen (including unknown) N = 1,106
Unknown	83	229
Viral subtype, n (%)		
B	196 (52%)	610 (55%)
C	47 (13%)	101 (9.1%)
CRF01_AE	29 (7.7%)	72 (6.5%)
CRF02_AG	34 (9.1%)	100 (9.0%)
F1	17 (4.5%)	53 (4.8%)
Other (n<50)	52 (14%)	170 (15%)

Table 2: Associations between non-integrase resistance mutations and individual and virological characteristics

Characteristic	All	3'PPT c9053t		Env Y61H		Env A539V		NC N8S	
	N = 1,106	Present N = 47	p-value[1]	Present N = 102	p-value[1]	Present N = 36	p-value[1]	Present N = 52	p-value[1]

Characteristic	All N = 1,106	3'PPT c9053t Present N = 47	p-value[1]	Env Y61H Present N = 102	p-value[1]	Env A539V Present N = 36	p-value[1]	NC N8S Present N = 52	p-value[1]
Age at ART start, Median (Q1, Q3)	32 (26, 42)	32 (26, 39)	0.16	30 (25, 44)	0.41	38 (31, 47)	**0.01**	30 (23, 37)	**0.02**
Gender, n (%)			0.22		0.07		0.41		0.84
Men (including trans men)	990 (90%)	45 (96%)		97 (95%)		31 (86%)		47 (90%)	
Women (including trans women)	116 (10%)	2 (4.3%)		5 (4.9%)		5 (14%)		5 (9.6%)	
Ethnic group, n (%)			0.37		0.96		**0.02**		0.38
White	806 (75%)	37 (80%)		75 (75%)		20 (57%)		41 (80%)	
All other ethnic groups	267 (25%)	9 (20%)		25 (25%)		15 (43%)		10 (20%)	
Probable route of HIV acquisition, n (%)			0.16		0.38		0.34		0.13
Sex between men	816 (78%)	39 (87%)		72 (75%)		25 (71%)		44 (86%)	
All other routes of HIV acquisition	230 (22%)	6 (13%)		24 (25%)		10 (29%)		7 (14%)	
Region of birth, n (%)			**0.03**		0.54		**0.004**		1.00
United Kingdom	665 (64%)	36 (80%)		62 (67%)		14 (41%)		33 (65%)	
Born outside the UK	367 (36%)	9 (20%)		30 (33%)		20 (59%)		18 (35%)	
Baseline CD4 (cells/µl), n (%)			0.96		0.07		0.67		0.16
<350 cells/µl	227 (21%)	10 (21%)		15 (15%)		9 (26%)		7 (13%)	
≥350 cells/µl	862 (79%)	37 (79%)		87 (85%)		26 (74%)		45 (87%)	
Baseline VL (copies/mL), n (%)			0.26		**0.04**		**0.04**		0.98
<100,000 copies/mL	394 (45%)	20 (51%)		44 (51%)		7 (23%)		20 (43%)	

Characteristic	All N = 1,106	3'PPT c9053t Present N = 47	p-value[1]	Env Y61H Present N = 102	p-value[1]	Env A539V Present N = 36	p-value[1]	NC N8S Present N = 52	p-value[1]
≥100,000 copies/mL	483 (55%)	19 (49%)		42 (49%)		23 (77%)		27 (57%)	
Viral subtype, n (%)			p<0.001		0.25		p<0.001		p<0.001
B	610 (55%)	40 (85%)		50 (49%)		5 (14%)		42 (81%)	
Non-B	495 (45%)	7 (15%)		52 (51%)		31 (86%)		10 (19%)	
Any INSTI SDRMs present, n (%)	8 (0.7%)	0 (0%)	1.00	1 (1.0%)	0.29	0 (0%)	1.00	2 (3.8%)	0.05
Any PI SDRMs present, n (%)	29 (2.7%)	1 (2.2%)	1.00	1 (1.0%)	0.50	1 (2.8%)	1.00	3 (5.9%)	0.18
Any NRTI SDRMs present, n (%)	44 (4.1%)	1 (2.1%)	0.72	2 (2.0%)	0.42	0 (0%)	0.40	2 (3.8%)	1.00
Any NNRTI SDRMs present, n (%)	35 (3.2%)	2 (4.3%)	0.67	2 (2.0%)	0.76	0 (0%)	0.63	0 (0%)	0.41
Any INSTI accessory mutations present, n (%)	85 (7.8%)	0 (0%)	0.07	1 (1.0%)	0.01	8 (22%)	0.002	2 (3.8%)	0.58

[1]Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

Table 3: Univariable and multivariable Cox regression associations between non-integrase mutations and outcomes

Characteristic	Viral suppression (n = 267)			Viral blip (n = 241)		
	HR	95% CI	p-value	HR	95% CI	p-value

Characteristic	Viral suppression (n = 267)			Viral blip (n = 241)		
	HR	95% CI	p-value	HR	95% CI	p-value
Unadjusted analysis						
3'PPT c9053t (unadjusted)						
Present	1.31	0.74, 2.31	0.3	1.44	0.52, 3.97	0.5
Env Y61H (unadjusted)						
Present	1.18	0.79, 1.74	0.4	0.66	0.28, 1.54	0.3
Env A539V (unadjusted)						
Present	0.88	0.49, 1.58	0.7	0.83	0.26, 2.67	0.8
NC N8S (unadjusted)						
Present	0.73	0.42, 1.26	0.3	0.57	0.18, 1.82	0.3
Adjusted analysis[1,2]						
3'PPT c9053t (adjusted)						
Present	1.29	0.71, 2.35	0.4	1.60	0.58, 4.46	0.4
Env Y61H (adjusted)						
Present	1.19	0.78, 1.81	0.4	0.75	0.32, 1.76	0.5
Env A539V (adjusted)						
Present	0.77	0.41, 1.42	0.4	0.96	0.30, 3.14	>0.9
NC N8S (adjusted)						
Present	0.59	0.34, 1.04	0.071	0.59	0.18, 1.90	0.4

[1]Analyses of viral suppression adjusted for age, gender, region of birth, baseline CD4 and baseline VL.
[2]Analyses of viral blip adjusted for age, gender, and baseline CD4.
Abbreviations: CI = Confidence Interval, HR = Hazard Ratio

Graphical Abstract