# The increasing genetic diversity of HIV-1 in the UK, 2002–2010

## The UK Collaborative Group on HIV Drug Resistance

**Objective:** HIV-1 is typically categorized by genetically distinct viral subtypes. Viral subtypes are usually compartmentalized by ethnicity and transmission group and, thus, convey important epidemiological information, as well as possibly influencing the rate of disease progression. We aim to describe the prevalence and time trends of subtypes observed among key populations living with HIV-1 in the UK.

**Design:** Analyses of reverse transcriptase and protease sequences generated from HIV-1-positive antiretroviral-naive patients as part of routine resistance testing between 2002 and 2010 in all public health and NHS laboratories in the UK.

**Methods:** Subtype was assigned centrally using the SCUEAL algorithm. Subtyping results were combined with data from the UK Collaborative HIV Cohort Study and the UK HIV and AIDS Reporting System. Analyses adjusted for the number of national HIV-1 diagnoses made each year within demographic subgroups. Viral subtypes were described overall, over time and by demographic subgroup.

**Results:** Subtype B diagnoses (39.9%) have remained stable since 2005, whereas subtype C diagnoses (34.3%) were found to decline in prevalence from 2004. Across most demographic subgroups, the prevalence of non-B non-C subtypes has increased over time, in particular novel recombinant forms (9.9%), subtype G (2.7%), and CRF01 AE (2.0%).

**Conclusion:** HIV-1 subtypes are increasingly represented across all demographic subgroups and this could be evidence of sexual mixing. Between 2002 and 2010, the prevalence of novel recombinant forms has increased in all demographic subgroups. This increasing genetic diversity and the effect of subtype on disease progression may impact future HIV-1 treatment and prevention.



*AIDS* 2014, **28**:773–780

**Keywords: epidemiology, genetic diversity, HIV-1, subtype, subtyping, UK**

## Introduction

Categorizing HIV-1 into genetic subtypes has been a powerful epidemiological tool to provide insights into patterns of HIV-1 transmission, with direct implications for public health policy. Genetically distinct viral strains are typically classified as one of nine distinct viral subtypes and mosaic viruses include 55 circulating recombinant forms (CRFs), unique recombinant forms (URFs) and other less well characterized complex structures [1,2]. Alongside epidemiological uses, genetically divergent strains have biological differences [3], which may impact on the development of drug resistance [4], the susceptibility and response to antiretroviral therapy (ART) [5] and the rate of disease progression [5–7], thus potentially bearing on clinical care.

The presence of non-B subtypes has been increasing in western Europe in recent years as a result of immigration from sub-Saharan Africa, Asia and eastern Europe [8]. This has led to multiple parallel epidemics within the UK, and subtypes are strongly associated with ethnicity and exposure group. In the UK, most HIV-1 group M subtypes and several CRFs are represented [9] and

Medical Research Council, London, UK.

Correspondence to Professor David Dunn, MRC Clinical Trials Unit at UCL, 125 Kingsway, Aviation House, London, WC2B 6NH, UK.

E-mail: d.dunn@ucl.ac.uk

Received: 2 July 2013; revised: 18 October 2013; accepted: 18 October 2013.

DOI:10.1097/QAD.0000000000000119

ISSN 0269-9370 

subtype B virus, typically associated with MSM, is the most prevalent. The subtypes observed in the UK have been recently described for MSM [10] and, historically, for heterosexuals [11]. We aim to describe the prevalence and time trends of subtypes observed among major demographic subgroups affected by HIV-1 in the UK.

## Methods

### Resistance and clinical data

The UK HIV Drug Resistance Database was established in 2001 and attempts to collect all resistance tests conducted within public health and NHS laboratories in the UK as part of routine clinical care. The resistance tests analysed in this study used bulk sequencing of the *pol* gene, encoding the protease and reverse transcriptase genes using a variety of in-house and commercial assays. Resistance tests conducted between 2002 and 2010 were used; from 2003 UK guidelines recommended that resistance testing be performed for all drug-naive patients prior to commencing treatment, and from 2005 [12] for all newly diagnosed patients.

Subtype was assigned centrally based on the Subtype Classification Using Evolutionary Algorithms (SCUEAL) [13], which uses model-based phylogenetic methods to assign viral subtype. Previous studies have used the REGA algorithm [14]; however, SCUEAL has been shown to leave a smaller number of sequences as unclassified, which are of growing importance [13]. Viruses with mosaic genomes which do not yet fulfil the criterion to be established as new CRFs were classified as novel recombinants. Sub-subtypes such as A1 and A2 or F1 and F2 were combined in our analysis, as were CRFs other than CRF01_AE, so that categories with small numbers could be avoided allowing changes over time in these groups to be analysed.

Demographic data were acquired via annual linkage (using pseudo-anonymized identifiers such as clinic number, soundex and date of birth) to the UK Collaborative HIV Cohort Study (UK CHIC), which includes patients from 13 of the largest clinics within the UK [15], and to the HIV and AIDS Reporting System (HARS) [16], coordinated by Public Health England. These data include all new diagnoses reported nationally since the beginning of the epidemic and data from a national cohort of persons in care updated annually since 1997.

### Analysis

The first ART-naive resistance sample for each patient in the UK HIV Drug Resistance Database conducted between 2002 and 2010 was analysed, in which year refers to the date of sample. Our analyses examined the following demographic subgroups: black African women, black African heterosexual men, white heterosexual men, white heterosexual women, black African MSM, white MSM and other MSM. Within each year and demographic subgroup, the proportion of resistance tests of each subtype was calculated. To adjust for the overall number of new diagnoses within the UK, these figures were then multiplied by the total number of new diagnoses [17] per year within each demographic subgroup to give the total number of diagnoses per year for each subtype. These temporal estimates were combined for an analysis of subtype by demographic subgroup and were combined across demographic subgroups for an analysis of overall temporal trends. More detailed analyses examined temporal trends by demographic subgroups and categorized a subtype as either subtype B, C or non-B, non-C to avoid categories with small numbers. Differences in subtype prevalence between groups were tested using the $\chi^2$ test. All analyses were conducted in Stata/IC 12.1 (StataCorp LP, College Station, Texas, USA).

## Results

### Study population

In total, 27 657 ART-naive resistance tests from 2002 to 2010 were available for analysis; this represents 46% of the 60 296 total diagnoses during this period. A total of 21 945 (79%) resistance tests had an associated gender, ethnicity and probable exposure group; 11 348 (52%) patients were MSM, 10 567 (48%) had a heterosexual exposure source and 30 had another exposure source; 11 784 (54%) patients were white, 7374 (34%) were black African, 847 (4%) were black Caribbean, 458 (2%) were of another black ethnicity and 1482 (7%) had another ethnicity. A diagnosis date was available for 20 159 (92%) patients. Approximately, two-thirds of the resistance tests were conducted within 3 months of diagnosis ($n = 13\,385$; 66%) and the median (interquartile range, IQR) time between diagnosis and resistance test date was 25 (7–394) days. The number of tests conducted within 3 months of diagnoses increased from 58% in 2002 to 75% in 2010, $P < 0.001$, reflecting national guidelines for resistance testing.

Demographic information was missing for 5712 resistance tests, but by adjusting for national diagnoses, any bias in the differential collection of demographic information will have been minimized. Patients with missing information were less likely to have subtype B virus ($n = 2708$; 47.4% compared with $n = 11\,354$; 51.7%; $P < 0.001$) and more likely to have subtype C virus ($n = 1628$; 28.5% compared with $n = 5714$; 26.0%; $P < 0.001$) than patients wherein this information was known. This suggests that probable exposure source information is more likely to be missing from transmission groups with a greater prevalence of

subtype C virus, such as the black African heterosexual population.

To determine how representative our data was to the UK epidemic, we compared sites in our study where more than one third of resistance tests conducted had demographic data to those that did not. As a result of the large sample size, significant differences were found between these sites. There were significantly more male patients in the sites with more comprehensive demographic data (61 vs. 59%; $P < 0.001$) and consequently there were also differences in the proportion of MSM (34 vs. 32%; $P < 0.001$) and black Africans (45 vs. 47%; $P < 0.001$). Overall, in absolute terms, these differences were felt to be small and unlikely to have a large impact on the prevalence of HIV-1 subtypes.

## Overall temporal trends

As is typical of other resource-rich countries, the predominant viral subtype was found to be subtype B ($n = 24\,040$; 39.9%). Other common subtypes were C ($n = 20\,678$; 34.3%), A ($n = 3006$; 5.0%), novel recombinants ($n = 5989$; 9.9%), other CRFs ($n = 1860$; 3.1%), D ($n = 1521$; 2.5%), G ($n = 1636$; 2.7%) and CRF01 AE ($n = 1224$; 2.0%). Table 1 presents the temporal trends for these main subtypes and Fig. 1 illustrates these. Major changes occurred between 2002 and 2005 as the estimated number of diagnoses per year increased from 2223 to 2855 for subtype B and for subtype C increased from 2372 to 2933 followed by a decrease to 2629. After 2006, the number of subtype B diagnoses remained relatively stable, whereas the number of subtype C diagnoses continued to decrease.

The number of novel recombinant diagnoses increased from 461 to 742 per year between 2002 and 2010; the constituent subtypes found in these recombinants are shown in Table 3 (Supplementary Digital Content, http://links.lww.com/QAD/A440) and these novel recombinants often contained fragments of subtypes G ($n = 575$), A ($n = 566$) and B ($n = 434$). Due to the overall decline in the number of yearly HIV-1 diagnoses, this increase meant that novel recombinants grew as a proportion of diagnosed infections and by 2010 accounted for 13.3% of all new diagnoses. CRF01_AE diagnoses increased from 49 to 173 per year between 2002 and 2006, and remained relatively stable after this. Other CRFs remained stable throughout the entire period studied and were mostly consisted of CRF02_AG ($n = 250$; 46% of other CRFs), CRF15_01B ($n = 77$; 14%) and CRF6_cpx ($n = 69$; 13%). Finally, the numbers of patients diagnosed with subtypes A and D declined disproportionately between 2002 and 2010, leading to a decrease in the overall proportion of infections with these subtypes.

## Subtype by demographic subgroup

The subtype diversity within demographic subgroups is shown in Table 2. MSM were found to predominantly,

**Table 1. Trends in subtype over time, 2002–2010.**

| Year | A | CRF01 AE | B | C | D | G | Other pure | CRFs | Novel recombinants | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| 2002 | 358; 5.9% (16) | 49; 0.8% (6) | 2223; 36.5% (437) | 2372; 39.0% (110) | 311; 5.1% (13) | 26; 0.4% (1) | 0; 0% (0) | 284; 4.7% (14) | 461; 7.6% (33) | 6084 (630) |
| 2003 | 485; 6.9% (34) | 74; 1.1% (8) | 2472; 35.1% (565) | 2731; 38.9% (166) | 364; 5.2% (20) | 153; 2.2% (10) | 28; 0.4% (4) | 191; 2.7% (15) | 530; 7.5% (53) | 7027 (875) |
| 2004 | 393; 5.3% (57) | 151; 2.1% (24) | 2672; 36.3% (982) | 2933; 39.9% (401) | 191; 2.6% (27) | 215; 2.9% (32) | 30; 0.4% (4) | 218; 3.0% (33) | 551; 7.5% (105) | 7355 (1665) |
| 2005 | 454; 6.0% (121) | 138; 1.8% (49) | 2855; 37.9% (1477) | 2629; 34.9% (707) | 158; 2.1% (42) | 253; 3.4% (72) | 40; 0.5% (13) | 210; 2.8% (59) | 789; 10.5% (245) | 7525 (2785) |
| 2006 | 282; 4.0% (113) | 173; 2.4% (56) | 2859; 40.5% (1577) | 2527; 35.8% (980) | 130; 1.8% (50) | 207; 2.9% (81) | 24; 0.3% (10) | 139; 2.0% (54) | 709; 10.1% (293) | 7052 (3214) |
| 2007 | 292; 4.2% (121) | 181; 2.6% (71) | 2903; 42.1% (1684) | 2190; 31.7% (896) | 122; 1.8% (51) | 212; 3.1% (88) | 34; 0.5% (15) | 200; 2.9% (88) | 770; 11.2% (346) | 6903 (3360) |
| 2008 | 288; 4.3% (130) | 156; 2.3% (64) | 2763; 41.2% (1730) | 2165; 32.3% (1009) | 107; 1.6% (47) | 267; 4.0% (124) | 33; 0.5% (16) | 216; 3.2% (101) | 704; 10.5% (353) | 6699 (3574) |
| 2009 | 215; 3.6% (103) | 168; 2.8% (64) | 2744; 45.6% (1532) | 1678; 27.8% (804) | 87; 1.4% (43) | 153; 2.5% (74) | 56; 0.9% (26) | 192; 3.2% (88) | 732; 12.1% (361) | 6024 (3095) |
| 2010 | 239; 4.3% (101) | 134; 2.4% (49) | 2550; 45.5% (1370) | 1453; 25.9% (641) | 50; 0.9% (21) | 150; 2.7% (68) | 69; 1.2% (34) | 213; 3.8% (103) | 742; 13.3% (360) | 5601 (2747) |
| Total | 3006; 5.0% (796) | 1224; 2.0% (391) | 24041; 39.9% (11 354) | 20 678; 34.3% (5714) | 1520; 2.5% (314) | 1636; 2.7% (550) | 314; 0.5% (122) | 1863; 3.1% (555) | 5988; 9.9% (2149) | 60 270 (21 945) |

Estimated total number of diagnoses; percentage of diagnoses per year (total number of resistance tests). CRF, circulating recombinant form.

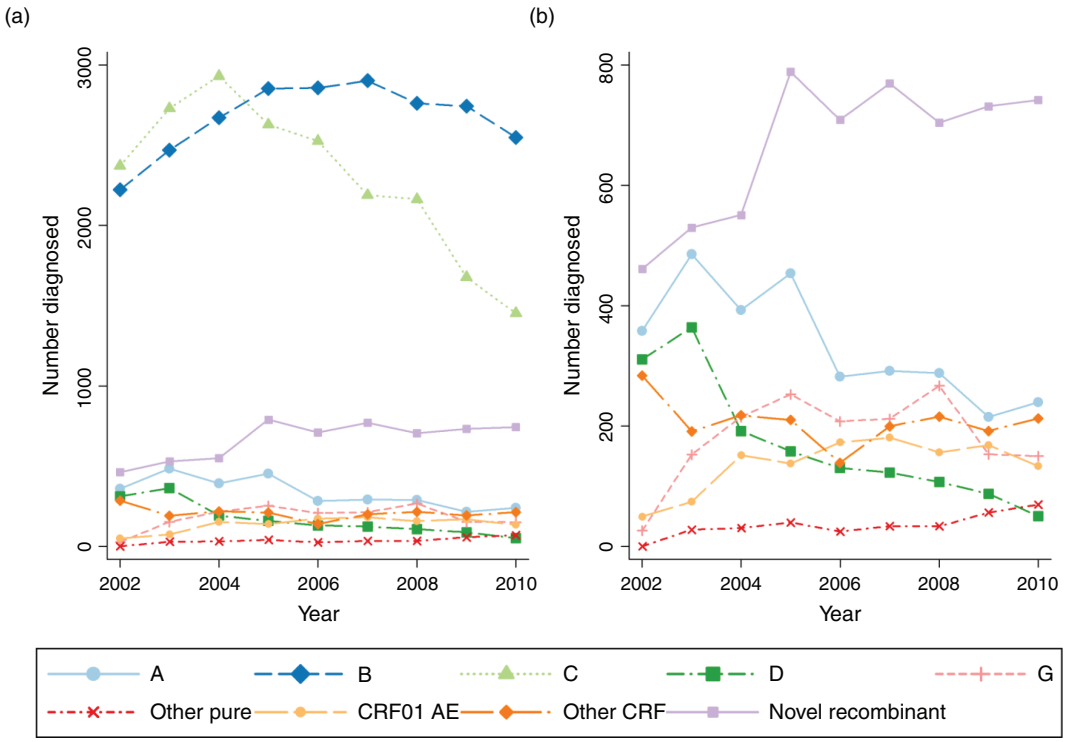


**Fig. 1. Trends in numbers of patients diagnosed by subtype between 2002 and 2010.** (a) All subtypes (b) non-B, non-C subtypes only (using different scale on *y*-axis).

though not exclusively, have subtype B virus ($n = 18\,847$; 88.4%). Although MSM had less subtype diversity than other transmission groups, all subtypes were represented and there were a notable number of patients infected with novel recombinant strains (1200; 5.6%). Patients with heterosexual contact as a probable exposure source had a much greater variety of subtypes, although there were gender differences in this group. The most frequent subtype for both genders was subtype C, although a greater proportion of women ($n = 12\,789$; 55.0%) have this subtype than men ($n = 7331$; 46.7%; $P < 0.001$). Subtype B infections were more common in heterosexual men ($n = 2947$; 18.8%) than women (2246; 9.7%; $P < 0.001$). Subtype A was also frequent in those with heterosexual contact as well as a variety of CRFs. Finally, novel recombinants were frequent in both men and women with a probable heterosexual exposure source and occurred in greater numbers within heterosexuals than MSM reflecting historical differences in viral diversity between these populations.

## Temporal trends by demographic subgroup

Temporal trends between 2002 and 2010 in subtype distribution by demographic subgroup are shown in Figs. 2 and 3 and Table 4 (Supplementary Digital Content, http://links.lww.com/QAD/A440) shows the subtype distribution among new diagnoses for heterosexual, black African or white, men and women. The total number of heterosexual black Africans diagnosed with HIV-1 fell from 2005/2006, which has similarly led to a decline in the total number of each subtype diagnosed. Trends were similar for black African men and

**Table 2. Subtype by probable exposure category, 2002–2010.**

| Subtype | Sex between men | Heterosexual contact | | Total |
|---|---|---|---|---|
| | | Men | Women | |
| A | 170; 0.8% (83) | 1013; 6.4% (235) | 1822; 7.8% (478) | 3006 (796) |
| B | 18 847; 88.4% (9975) | 2947; 18.8% (692) | 2246; 9.7% (687) | 24 040 (11 354) |
| C | 558; 2.6% (305) | 7331; 46.7% (1884) | 12 789; 55.0% (3525) | 20 678 (5714) |
| D | 38; 0.2% (17) | 501; 3.2% (102) | 983; 4.2% (195) | 1521 (314) |
| G | 70; 0.3% (37) | 570; 3.6% (188) | 995; 4.3% (325) | 1636 (550) |
| Other pure | 71; 0.3% (44) | 85; 0.5% (27) | 157; 0.7% (51) | 314 (122) |
| CRF01_AE | 151; 0.7% (95) | 806; 5.1% (187) | 267; 1.2% (109) | 1224 (391) |
| Other CRF | 215; 1.0% (122) | 644; 4.1% (174) | 1001; 4.3% (259) | 1860 (555) |
| Novel recombinants | 1200; 5.6% (666) | 1816; 11.6% (534) | 2973; 12.8% (949) | 5989 (2149) |
| Total | 21 321; 35.4% (11 344) | 15 714; 26.1% (4023) | 23 234; 38.5% (6578) | 60 270 (21 945) |

Estimated total number of diagnoses; percentage of exposure category (total number of resistance tests). CRF, circulating recombinant form.

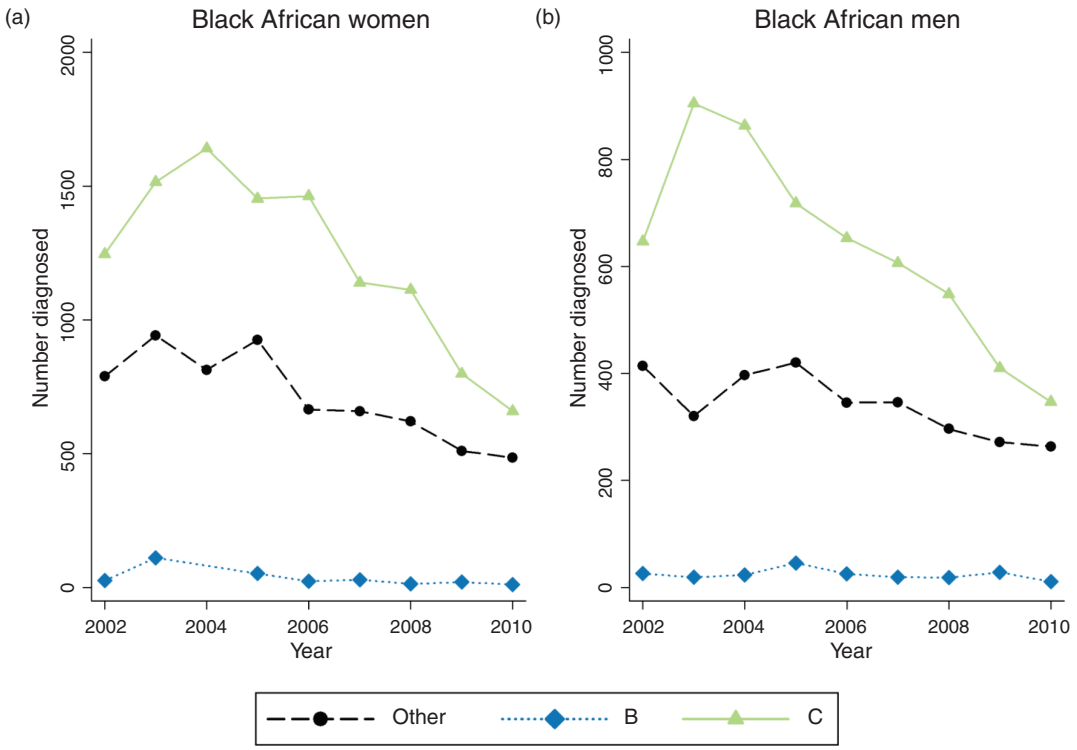


**Fig. 2. Trends in subtype over time among black African heterosexuals, 2002–2009.** Note different scales on *y*-axis for women and men.

women with an increase in the proportion of patients diagnosed with non-B non-C subtypes and a decrease in the proportion of subtype C. Non-B non-C subtypes were diagnosed in 485 and 264 black African women and men, respectively, in 2010. The declining trend was most prominent for subtypes A, D and other CRFs but was partially offset by an increase in the absolute number and proportion infected with novel recombinant virus. For white heterosexual men and women, there are no clear time trends, although with limited data available for

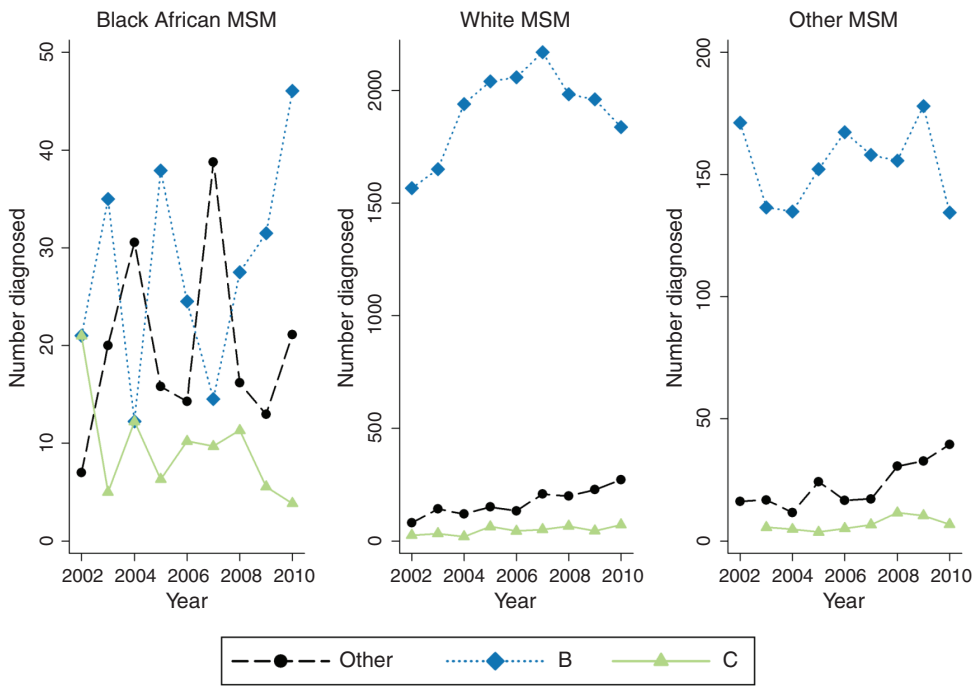


**Fig. 3. Trends in subtype over time among MSM, 2002–2009.** Note different scales on *y*-axis by ethnic group.

2002–2004. White heterosexual women are roughly equally divided between subtype B, C and non-B non-C, whereas white heterosexual men have proportionately more subtype B infections (41.1% of overall infections) and proportionately fewer subtype C infections (22.5% of overall infections) compared with women ($P<0.001$).

The number of HIV-1 infections diagnosed among MSM increased prior to 2007 and has remained high since then in all ethnic groups. This has been accompanied by an increase in the total number of non-B non-C subtypes diagnosed during this period. Substantial differences in the subtype make-up of diagnosed MSM by ethnicity are evident. White MSM and other MSM were chiefly diagnosed with subtype B, whereas black African MSM were estimated to have a mixture of subtype B (250; 48.9%), C (85; 16.7%) and non-B, non-C (177; 34.6%). However, the number and proportion of non-B, non-C subtypes for white MSM increased markedly from 80 (4.8%) in 2002 to 272 (12.4%) in 2010. This increase was largely driven by novel recombinant viruses, accounting for 8.0% of diagnoses among white MSM in 2010. Trends cannot be discerned for other ethnicities due to the small numbers diagnosed.

## Discussion

This study found broad representation of HIV-1 subtypes within the UK, both overall and within specific demographic subgroups. Subtype should no longer be used as a proxy for transmission group when this is unknown, as has been the case historically, as heterosexual infections now account for 22% of all subtype B infections. The increase in non-B non-C subtypes, particularly novel recombinants among MSM of white and other ethnicity, demonstrates that there is ongoing emergence of new novel recombinant forms in the UK or that these variants are more transmissible. The number of patients diagnosed with subtype B virus has remained stable and is the most prevalent subtype in the UK, whereas the annual number of subtype C HIV-1 diagnoses, has fallen since 2004, reflecting a decline in new HIV-1 diagnoses in the black African population. However, it remains the second most prevalent subtype overall and remains the most common subtype in the black African population.

Previous UK studies [18] have also found 'an increasingly broad range of genetic diversity' and using phylogenetic surveillance showed that subtypes typically associated with infections acquired in East Africa were spreading among MSM in the UK. Gifford *et al.* [18] used 10 537 sequences collected between 1996 and 2004, so our study represents a significant update on the changes in the prevalence of subtypes that have occurred since then, while also focusing on temporal trends within probable exposure groups in greater detail. Another UK study [10] examined the prevalence of non-B subtypes among MSM diagnosed from 1980 to 2007, and also found that nonsubtype B infections were increasing in the UK and suggested that the association of subtypes to socio-demographic groups was weakening.

A recent meta-analysis [19] of global data, with 65 913 samples, has concluded that there has been an increase in the proportion of the subtypes A, F, G, H, CRF01 AE and CRF02 AG and that the global proportion of all CRFs has also grown, but did not describe this by demographic subgroup. However, URFs decreased as a proportion of infections by 3.1% between 2000–2003 and 2004–2007. Abecasis *et al.* [20] used data from the SPREAD European surveillance programme in 2002–2005 and found that subtype was highly determined by demographic subgroups and suggested highly compartmentalized, parallel, epidemics but only CRF02_AG and subtype F had significant time trends. Our more recent data suggest that the effect of increased migration to the UK has led to a very different epidemic to the rest of Europe, although the compartmentalization found in SPREAD may have also broken down since 2005. Other studies [21–23] have also observed increasing genetic diversity demonstrating that this is not a phenomenon unique to the UK.

Our study population comprises patients who received a drug resistance test and may not therefore be representative of the entire HIV-1-positive population in the UK. We tried to minimize this bias by only using tests conducted after 2001 and by accounting for differential use of resistance testing by ethnicity and exposure group. However, we could not adjust for the 24% of the HIV-1-positive population estimated to not yet be aware of their infection [24]. As a consequence, we may be under-reporting some non-B subtypes as previous research has shown that non-African born heterosexuals have a higher undiagnosed rate of infection compared with MSM (31 compared with 20% [24]). For our temporal trend analyses, we used the year of sampling for the resistance test, which might not reflect the date of seroconversion, although this should not change our main conclusions. It would have been interesting to have conducted further analyses comparing UK acquired infections to those acquired abroad. An indirect method for estimating country of infection, based on year of migration and $CD4^+$ cell count at diagnosis, was recently published [25] and showed that 33% of heterosexuals born abroad acquired their infection in the UK. However, a similar analysis was not possible in this study due to incomplete data on these two variables.

The choice of SCUEAL as the automated subtyping tool could have influenced our findings compared to other studies, which have generally used the REGA v2 algorithm. A comparison between the subtype assigned by SCUEAL and REGA v2 to sequences in this study is

shown in Table 6 (Supplementary Digital Content, http://links.lww.com/QAD/A440). Consistent with another analysis [26], there was strong concordance between the pure subtypes; the main difference between the algorithms is that viruses classified as CRF02_AG by REGA are usually classified as either G, A/G recombinants with alternate breakpoints, or as a complex subtype, by SCUEAL. The estimated prevalence of these subtypes would, therefore, have been affected accordingly. Our analysis uses partial sequences of the *pol* gene, comprising approximately 10% of the viral genome, as this is the region typically sequenced for resistance testing. This results in underestimates of the prevalence of novel recombinant forms as mosaic viruses may have breakpoints outside of this region. This bias cannot currently be quantified, but the increasing availability of full length sequences from next generation sequencing platforms will allow future quantification of this effect.

The increasing diversity of HIV-1 could be viewed as evidence that any fitness differences between subtypes are likely to be small. The fact that distinct demographic subgroups have overlapping subtypes suggests sexual linkage between different subgroups, although this can only be confirmed by conducting further phylogenetic analyses. The increase in novel recombinants is further evidence that sexual linkage between demographic subgroups is occurring for individuals to be simultaneously infected with genetically distinct viruses. The impact this increase in novel recombinants will have on disease progression, treatment and the development of drug resistance is not yet known.

## Acknowledgements

The UK Collaborative Group on HIV Drug Resistance is a collaboration between the UK HIV Drug Resistance Database; UK CHIC; Public Health England HARS; and participating academic centres, clinics, and laboratories.

Members of Analysis/Writing Group are as follows: David Dolling, Stéphane Hué, Valerie Delpech, Esther Fearnhill, Andrew Leigh Brown, Anna Maria Geretti, Deenan Pillay, David Dunn.

Members of Steering Committee are as follows: Celia Aitken, Gartnavel General Hospital, Glasgow; David Asboe, Anton Pozniak, Chelsea & Westminster Hospital, London; Daniel Webster, Royal Free NHS Trust, London; Patricia Cane, Health Protection Agency, Porton Down; Hannah Castro, David Dunn (Co-Chair), David Dolling, Esther Fearnhill, Anna Tostevin, Kholoud Porter, MRC Clinical Trials Unit, London; David Chadwick, South Tees Hospitals NHS Trust, Middlesbrough; Duncan Churchill, Brighton and Sussex University Hospitals NHS Trust; Duncan Clark, St Bartholomew's and The London NHS Trust; Simon Collins, HIV i-Base, London; Valerie Delpech, Health Protection Agency, Centre for Infections, London; Anna Maria Geretti, Institute of Infection and Global Health, University of Liverpool; David Goldberg, Health Protection Scotland, Glasgow; Antony Hale, Leeds Teaching Hospitals NHS Trust; Stéphane Hué, University College London; Steve Kaye, Imperial College London; Paul Kellam, Wellcome Trust Sanger Institute & UCL Medical School; Linda Lazarus, Expert Advisory Group on AIDS Secretariat, Health Protection Agency, London; Andrew Leigh-Brown, University of Edinburgh; Nicola Mackie, Imperial NHS Trust; Chloe Orkin, St. Bartholomew's Hospital, London; Philip Rice, St George's Healthcare Trust, London; Deenan Pillay (Co-Chair), Andrew Phillips, Caroline Sabin, University College London Medical School; Erasmus Smit, Health Protection Agency, Birmingham Heartlands Hospital; Kate Templeton, Royal Infirmary of Edinburgh; Peter Tilston, Manchester Royal Infirmary; William Tong, Guy's and St. Thomas' NHS Foundation Trust, London; Ian Williams, Mortimer Market Centre, London; Hongyi Zhang, Addenbrooke's Hospital, Cambridge; Mark Zuckerman, King's College Hospital, London.

Centres contributing data: Clinical Microbiology and Public Health Laboratory, Addenbrooke's Hospital, Cambridge (Jane Greatorex); HIV/GUM Research Laboratory, Chelsea and Westminster Hospital, London (Adrian Wildfire); Guy's and St. Thomas' NHS Foundation Trust, London (Siobhan O'Shea, Jane Mullen); HPA – Public Health Laboratory, Birmingham Heartlands Hospital, Birmingham (Erasmus Smit); HPA London (Tamyo Mbisa); Imperial College Health NHS Trust, London (Alison Cox); King's College Hospital, London (Richard Tandy); Medical Microbiology Laboratory, Leeds Teaching Hospitals NHS Trust (Tony Hale, Tracy Fawcett); Specialist Virology Centre, Liverpool (Mark Hopkins, Lynn Ashton); Department of Clinical Virology, Manchester Royal Infirmary, Manchester (Peter Tilston); Department of Virology, Royal Free Hospital, London (Claire Booth, Ana Garcia-Diaz); Edinburgh Specialist Virology Centre, Royal Infirmary of Edinburgh (Jill Shepherd); Department of Infection & Tropical Medicine, Royal Victoria Infirmary, Newcastle (Matthias L Schmid, Brendan Payne); South Tees Hospitals NHS Trust, Middlesbrough (David Chadwick); St George's Hospital, London (Phillip Hay, Phillip Rice, Mary Paynter); Department of Virology, St Bartholomew's and The London NHS Trust (Duncan Clark, David Bibby); Molecular Diagnostic Unit, Imperial College, London (Steve Kaye); University College London Hospitals (Stuart Kirk); West of Scotland Specialist Virology Lab Gartnavel, Glasgow (Alasdair MacLean, Celia Aitken, Rory Gunson).

Coordinating Centre: Medical Research Council Clinical Trials Unit (MRC CTU), London (Kate

Coughlin, David Dolling, David Dunn, Esther Fearnhill, Anna Tostevin, Lorraine Fradette, Kholoud Porter).

This work was supported by the UK Medical Research Council (grant G0900274) and the European Community's 7th framework programme (FP7/2007–2013) under the Collaborative HIV and Anti-HIV Drug Resistance Network (CHAIN; project 223131).

## Conflicts of interest

None of the authors have an association that might pose a conflict of interest for this piece of work.

## References

1. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, *et al*. **HIV-1 nomenclature proposal.** *Science* 2000; **288**:55–56.
2. Los Alamos National Laboratory. HIV Circulating Recombinant Forms (CRFs). http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html [Accessed 28 August 2013].
3. Geretti AM, Harrison L, Green H, Sabin C, Hill T, Fearnhill E, *et al*. **Effect of HIV-1 subtype on virologic and immunologic response to starting highly active antiretroviral therapy.** *Clin Infect Dis* 2009; **48**:1296–1305.
4. Holguin A, Ramirez de Arellano E, Rivas P, Soriano V. **Efficacy of antiretroviral therapy in individuals infected with HIV-1 non-B subtypes.** *AIDS Rev* 2006; **8**:98–107.
5. Pai NP, Shivkumar S, Cajas JM. **Does genetic diversity of HIV-1 non-B subtypes differentially impact disease progression in treatment-naive HIV-1-infected individuals? A systematic review of evidence: 1996–2010.** *JAIDS* 2012; **59**:382–388.
6. Klein M, Odueyungbo A, Scherrer A, Ledergerber B, Fearnhill E, Sabin C, *et al*. Impact of viral subtype on immunologic and clinical disease progression in antiretroviral-naïve HIV-infected adults. *18th Conference on Retroviruses and Opportunistic Infections*. Boston, USA; 2011.
7. Lodi S, Phillips A, Touloumi G, Geskus R, Meyer L, Thiebaut R, *et al*. **Time from human immunodeficiency virus seroconversion to reaching CD4+ cell count thresholds < 200, <350, and <500 cells/mm(3): assessment of need following changes in treatment guidelines.** *Clin Infect Dis* 2011; **53**:817–825.
8. Thomson MM, Najera R. **Increasing HIV-1 genetic diversity in Europe.** *J Infect Dis* 2007; **196**:1120–1124.
9. Gifford R, de Oliveira T, Rambaut A, Myers RE, Gale CV, Dunn D, *et al*. **Assessment of automated genotyping protocols as tools for surveillance of HIV-1 genetic diversity.** *AIDS* 2006; **20**: 1521–1529.
10. Fox J, Castro H, Kaye S, McClure M, Weber JN, Fidler S. **Epidemiology of non-B clade forms of HIV-1 in men who have sex with men in the UK.** *AIDS* 2010; **24**:2397–2401.
11. Tatt ID, Barlow KL, Clewley JP, Gill ON, Parry JV. **Surveillance of HIV-1 subtypes among heterosexuals in England and Wales, 1997–2000.** *J Acquir Immune Defic Syndr* 2004; **36**:1092–1099.
12. Gazzard B, Comm BW. **British HIV Association (BHIVA) guidelines for the treatment of HIV-infected adults with antiretroviral therapy (2005).** *HIV Med* 2005; **6**:1–61.
13. Kosakovsky Pond SL, Posada D, Stawiski E, Chappey C, Poon AF, Hughes G, *et al*. **An evolutionary model-based algorithm for accurate phylogenetic breakpoint mapping and subtype prediction in HIV-1.** *PLoS Comput Biol* 2009; **5**:e1000581.
14. de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, *et al*. **An automated genotyping system for analysis of HIV-1 and other microbial sequences.** *Bioinformatics* 2005; **21**:3797–3800.
15. The UK Collaborative HIV Cohort Steering Committee. **The creation of a large UK-based multicentre cohort of HIV-infected individuals: the UK Collaborative HIV Cohort (UK CHIC) Study.** *HIV Med* 2004; **5**:115–124.
16. Public Health England Centre for Infections. The HIV and AIDS Reporting System (HARS). http://www.hpa.org.uk/webw/HPAweb&HPAwebStandard/Page/1317134347993. [Accessed 23 May 2013].
17. Public Health England Centre for Infections. United Kingdom New HIV Diagnoses to end of June 2012. http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/HIV/NewHIVDiagnoses/. [Accessed 23 May 2013].
18. Gifford RJ, de Oliveira T, Rambaut A, Pybus OG, Dunn D, Vandamme AM, *et al*. **Phylogenetic surveillance of viral genetic diversity and the evolving molecular epidemiology of human immunodeficiency virus type 1.** *J Virol* 2007; **81**:13050–13056.
19. Hemelaar J, Gouws E, Ghys PD, Osmanov S. **Global trends in molecular epidemiology of HIV-1 during 2000–2007.** *AIDS* 2011; **25**:679–689.
20. Abecasis AB, Wensing AM, Paraskevis D, Vercauteren J, Theys K, Van de Vijver DA, *et al*. **HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics.** *Retrovirology* 2013; **10**:7.
21. Chibo D, Birch C. **Increasing diversity of human immunodeficiency virus type 1 subtypes circulating in Australia.** *AIDS Res Hum Retroviruses* 2012; **28**:578–583.
22. Chen JH, Wong KH, Chen Z, Chan K, Lam HY, To SW, *et al*. **Increased genetic diversity of HIV-1 circulating in Hong Kong.** *Plos One* 2010; **5**:e12198.
23. Torimiro JN, D'Arrigo R, Takou D, Nanfack A, Pizzi D, Ngong I, *et al*. **Human immunodeficiency virus type 1 intersubtype recombinants predominate in the AIDS epidemic in Cameroon.** *New Microbiol* 2009; **32**:325–331.
24. Health Protection Agency. *HIV in the United Kingdom: 2012 Report*. Edited by Health Protection Agency; 2012.
25. Rice BD, Elford J, Yin Z, Delpech VC. **A new method to assign country of HIV infection among heterosexuals born abroad and diagnosed with HIV.** *AIDS* 2012; **26**:1961–1966.
26. Pineda-Peña AC, Faria NR, Imbrechts S, Libin P, Abecasis AB, Deforche K, *et al*. **Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: performance evaluation of the new REGA version 3 and seven other tools.** *Infect Genet Evol* 2013; **19**:337–348.