Epidemiol. Infect. (2004), **132**, 693–698. © 2004 Cambridge University Press
DOI: 10.1017/S095026880400233X Printed in the United Kingdom

HIV-1 subtype in Scotland: the establishment of a national surveillance system

D. L. YIRRELL[1,4,5]*, L. SHAW[2], S. M. BURNS[3], S. O. CAMERON[1], M. QUIGG[4], E. CAMPBELL[1] AND D. GOLDBERG[2,5]

[1] *Specialist Virology Centre, Gartnavel General Hospital, Glasgow, Scotland, UK*
[2] *Scottish Centre for Infection and Environmental Health, Glasgow, Scotland, UK*
[3] *Regional Clinical Virus Laboratory, Edinburgh Royal Infirmary, Edinburgh, Scotland, UK*
[4] *Centre for HIV Research, University of Edinburgh, Scotland, UK*
[5] *Department of Public Health, University of Glasgow, Scotland, UK*

(*Accepted 26 February 2004*)

SUMMARY

Historically, subtype B viruses in men who have sex with men (MSM) and injecting drug users (IDU) dominated the HIV epidemic in the United Kingdom, whereas non-B heterosexual infections dominate globally. Heterosexual contact is now the most common route of transmission in the United Kingdom. Here we monitor HIV subtype in Scotland, and link it to origin of infection. HIV-1 sequence was generated from new diagnoses and the subtype thus obtained linked with demographic data. Virus was subtyped from 80% (137/171) of all new diagnoses in Scotland. Of 58 individuals infected by heterosexual contact, 74% (43) harboured non-B viruses, contrasting with 7% (5/68) of those infected by IDU or MSM. Eighty-four per cent of non-Bs (46/55) were probably acquired outside the United Kingdom, but nine individuals probably acquired their non-B infection in the United Kingdom. Non-B subtypes of HIV-1 predominate in recently diagnosed, heterosexually acquired infections in Scotland and are present in all risk groups, even those with no exposure outside the United Kingdom.

INTRODUCTION

Members of the HIV-1 M group are responsible for the vast majority of infections leading to AIDS across the world. This group is subdivided into subtypes on the basis of sequence. At present there are nine subtypes designated A–D, F–H, J and K, which differ from each other within envelope gene amino-acid sequence by approximately 25–35% [1, 2]. In addition there are circulating recombinant forms (CRF), in which different parts of the genome appear to have been derived from different subtypes. These are distinct from recombinants that may arise in any one individual and remain unique to that patient, in that their mosaic structure has remained stable following serial transmission within a community. Some CRFs had previously been classified as one of the alphabet letters missing from the above list but, with the generation of more sequence data, have been re-categorized. One such example is the subtype E virus, common in Thailand which is now termed CRF01_AE, since its genome is composed of parts of subtypes A and E.

Subtypes of HIV-1 are not evenly distributed around the world. A single subtype tends to dominate a geographical region and/or risk group. In the United States and western Europe the epidemic

* Author for correspondence: Dr D. L. Yirrell, West of Scotland Specialist Virology Centre, Gartnavel General Hospital, 1053 Great Western Road, Glasgow G12 0YA, Scotland, UK.

of HIV was established, almost exclusively, by subtype B viruses and, in the main, was restricted to men who have sex with men (MSM) and injecting drug users (IDU) [3, 4]. By contrast, the main route of transmission in the rest of the world, is unprotected sexual intercourse among heterosexuals. This is particularly evident in sub-Saharan Africa, where the disease is thought to have originated, where 70% of worldwide infections have occurred, and where all HIV subtypes are found. The prevalence of subtype B infection in this region, however, is extremely low. Such observations led to the hypothesis that the efficiency of transmission varies among different subtypes and, in particular, subtype B viruses are less able than non-B viruses to transmit via the heterosexual route [5]. Although such theories cannot be supported by experimental evidence, the concern remains that non-B subtypes might drive new epidemics of HIV among heterosexuals in western countries.

In Scotland, during the last 4 years, more new HIV diagnoses have occurred in heterosexuals with exposure in Africa and Asia than in any other risk groups. Previously, most diagnoses had been in IDUs and MSM. If non-B subtypes were being transmitted from those with exposure in Africa and Asia to indigenous heterosexuals without such exposure, this dynamic could be identified through a surveillance programme which incorporated HIV subtype testing. Such information would be vital to public health professionals responsible for HIV prevention initiatives.

Additionally, the subtyping of HIV among new diagnoses may provide information of use to the clinician. Evidence exists for a more aggressive disease caused by subtype D compared to subtype A [6], an observation which may be extrapolated to suggest a subtype association with rates of disease progression, supported by a higher frequency of syncytium-inducing viruses in subtype D compared to subtype C [7]. Subtypes may also have different sensitivity to anti-retroviral therapy [8] and are known to have differential quantitation in some viral load assays [9].

Accordingly, Scotland's two specialist HIV testing laboratories in Edinburgh and Glasgow, and its national infection surveillance unit, the Scottish Centre for Infection and Environmental Health (SCIEH), were funded in 2000 to subtype infections among all new diagnoses. Here we report the findings of the initiative's first year.

METHODS

Design

The design is a subtype analysis of HIV-1 infection among newly diagnosed individuals in Scotland and the alignment of the resulting data with corresponding demographic and risk information.

Patients and samples

Virus from all new HIV-1 diagnoses, confirmed by Scotland's two specialists testing laboratories, was eligible for subtype analysis. Residual sera or plasma specimens submitted for HIV diagnostic or viral load testing respectively, were used as source material. These were stored at $-20\,^{\circ}C$, and processed in batches at 2–3 monthly intervals. The batches included HIV-negative and HIV-positive (subtype known) controls for internal quality assurance purposes.

Laboratory methods

RNA, extracted from either plasma or serum, was subjected to reverse transcription and PCR (RT–PCR), and amplified by nested PCR in both the v3/v4 region of the *env* gene (approximately 400 bp) and an approximately 720-bp region of the *gag* gene encompassing the p17/p24 junction [10]. PCR products were then sequenced on either an ABI 377 or 373 automated sequencer. Sequences thus obtained were aligned with homologous regions from reference viral strains obtained from the Los Alamos Database [11], using the Genetic Data Environment and BioEdit packages [12, 13]. Neighbour joining phylogenetic trees were constructed using the TREECON package [14], employing a Kimura distance matrix and each virus assigned a subtype by association of its sequence with a reference strain.

Patient demographic and risk information

Since 1985 HIV-testing laboratories throughout Scotland have reported cases of HIV seropositivity to SCIEH. The information held on each case includes gender, date of birth, soundex code of surname (a three-figure coding system which allocates a number to consonants and ignores vowels), first part of the postal code of residence, date of earliest positive specimen and risk exposure. Data are collected through the use of a national HIV test form. For cases whose risk status, as recorded on request forms by

https://doi.org/10.1017/S095026880400233X Published online by Cambridge University Press

the attending clinician, is IDU and/or homosexual/bisexual male, no active surveillance to verify this information is undertaken. For cases whose risk status is recorded as either child of infected mother or blood/blood product recipient, verification is conducted through liaison with the United Kingdom's Institute of Child Health (which coordinates UK-wide, clinician-based, HIV paediatric and obstetric surveillance systems) and the Scottish National Blood Transfusion Service respectively. Where heterosexual exposure is the only risk indicated or where incomplete information is provided, contact is made with the patient's physician. For the former group, the enquiry seeks to elucidate the risk status of the case's sexual partner(s) (e.g. IDU), in which part of the world the likely exposure occurred and, if this was outside the United Kingdom, whether or not the case was an indigenous Scot who travelled abroad on vacation/business.

RESULTS

During the first year of this surveillance initiative, 171 persons in Scotland were diagnosed HIV-1 seropositive. Specimens from 89% (153/171) were obtained by the investigators for subtype analysis. Sequence from the env and/or gag region of HIV was generated from 90% (137/153) of these. Accordingly, data on 80% of all infections newly diagnosed in Scotland between April 2000 and April 2001 are presented.

For each of the 137 cases, details of HIV subtype, presumed risk-group status and geographical region of exposure are presented in the Table. Overall, 40% (55/137) of cases harboured a non-B subtype infection. Of these, the majority (62%; 34/55) were subtype C, although subtypes A (nine), A/G [CRF02_AG] (six), E [CRF01_AE] (three) and G (three), were also found. Seventy-five per cent (46/61) of cases with a history of non-UK exposure harboured a non-subtype B virus, whereas this figure was only 12% (9/76) for those with UK-only exposure.

MSM/IDU

Of the 68 cases belonging to MSM (58) and IDU (10) groups, 7% (5/68) had a non-B subtype. For three of the five cases (one IDU, two MSM – all subtype C), no history of risk exposure outside the United Kingdom was evident.

Heterosexual men and women

Of the non-IDU heterosexual cases, 74% (43/58) harboured non-B subtypes, of which two thirds (29/43; 67%) were subtype C. Of cases with a non-UK exposure, mainly in Africa, 90% (38/42) harboured a non-B subtype. In contrast, 31% (5/16) of those with UK-only exposure belonged to this category.

Other risk

Seven of the 11 cases in this category were infected with non-B subtypes, three of which were infected perinatally. Of these, two (both subtype A) were born to African mothers, and one, with subtype A/G virus was born to a mother who had heterosexual exposure only in the United Kingdom.

Specimens without subtype data

No subtype is recorded for 34 specimens. This was either because they failed to amplify (16) or no sample was available (18). The 18 samples, which were not available consisted of two IDU and five MSM (all of whom reported their only exposure within the United Kingdom), six heterosexuals (five females; two reporting UK-only exposure and three African exposure and one male reporting African exposure), four with an unknown route of infection but probably within the United Kingdom and one person who received a blood transfusion in Asia. The 16 which failed to amplify, consisted of seven heterosexuals (three UK-only, two Africa and one USA/Caribbean exposure), two IDU (one UK-only, one European exposure) and eight MSM (seven UK-only, one European exposure).

DISCUSSION

The principal aim of establishing an HIV-1 subtype surveillance system in Scotland was to determine if non-B subtypes are being observed among individuals for whom there is no evidence of risk exposure outside the United Kingdom. In this context, data for the first year of the initiative indicated non-B infection in nine individuals – five heterosexuals, two men who had had sex with other men, one IDU and one perinatally infected child. In interpreting these results, recognition should be given to possible inaccuracies in the data as a consequence of methodological weaknesses. First, new diagnoses of HIV-1 do not

https://doi.org/10.1017/S095026880400233X Published online by Cambridge University Press

Table. *Subtypes of persons diagnosed HIV-1 positive in Scotland during 1 April 2000 to 1 April 2001 by risk group and region of exposure*

Risk group		Region of exposure	Subgroup							Total typed	% Non-B	Failure to type	No. spec. available	Total no. of diagnoses
			A	A/G	C	E	G	Non-B	B					
Heterosexual	UK		—	2	2	—	1	5	11	16	31	3	2	21
	Non-UK	Total	5	1	27	3	2	38	4	42	90	3	4	49
		Africa	5	1	26	—	2	34	1	35	97	2	4	41
		Americas/Carib/Australasia	—	—	—	—	—	—	—	—	—	1	—	1
		Asia	—	—	1	3	—	4	—	4	100	—	—	4
		Europe	—	—	—	—	—	—	3	3	0	—	—	3
	Total		5	3	29	3	3	43	15	58	74	6	6	70
IDU	UK		—	—	1	—	—	1	5	6	16	1	2	9
	Non-UK	Total	—	1	—	—	—	1	3	4	25	1	—	5
		Africa	—	1	—	—	—	1	—	1	100	—	—	1
		Americas/Carib/Australasia	—	—	—	—	—	—	1	1	0	—	—	1
		Asia	—	—	—	—	—	—	—	—	—	—	—	—
		Europe	—	—	—	—	—	—	2	2	0	1	—	3
	Total		—	1	1	—	—	2	8	10	20	2	2	14
MSM	UK		—	—	2	—	—	2	49	51	3	7	5	63
	Non-UK	Total	1	—	—	—	—	1	6	7	14	1	—	8
		Africa	—	—	—	—	—	—	—	—	—	—	—	—
		Americas/Carib/Australasia	—	—	—	—	—	—	5	5	0	—	—	5
		Asia	—	—	—	—	—	—	—	—	—	—	—	—
		Europe	1	—	—	—	—	1	1	2	50	1	—	3
	Total		1	—	2	—	—	3	55	58	5	8	5	71
Other	UK		—	1	—	—	—	1	2	3	33	—	4	7
	Non-UK	Total	3	1	2	—	—	6	2	8	20	—	1	9
		Africa	3	1	1	—	—	5	—	5	100	—	—	5
		Americas/Carib/Australasia	—	—	—	—	—	—	—	—	—	—	—	—
		Asia	—	—	—	—	—	—	1	1	0	—	1	2
		Europe	—	—	1	—	—	1	1	2	50	—	—	2
	Total		3	2	2	—	—	7	4	11	64	—	5	16
All	UK		—	3	5	—	1	9	67	76	12	11	13	100
	Non-UK		9	3	29	3	2	46	15	61	75	5	5	71
		Total	9	6	34	3	3	55	82	137	40	16	18	171

MSM, Men who have sex with men (includes 5 men who have sex with men and inject drugs).
IDU, Injecting drug user.
Other includes 5 blood/tissue recipients, 3 mother-to-child transmissions, and 8 cases where risk has not been established.

necessarily indicate new infections since some individuals may have been unaware of their status for some time and others may have been diagnosed in other countries. Secondly, cases are eligible for inclusion into the non-UK exposure category only if there is evidence of non-UK exposure. If there is none, cases default to the UK-only group. Information about geographical exposure is gleaned from cases' physicians, usually during the 12-month period following HIV diagnosis. The group of patients categorized with the greatest certainty, are immigrants/asylum seekers/students from countries, particularly those in sub-Saharan Africa, where HIV prevalence is high, as their status is usually self-evident to a physician. Apart from these cases, the physician's ability to provide SCIEH with information about the patients' risk status and likely geographical region of exposure is dependent on whether or not they are prepared to divulge accurate and comprehensive behavioural/demographic details. Seventeen years of HIV data-collection experience at SCIEH indicates that most data are robust. Indeed, the strong, anticipated associations between geographical region of exposure and HIV subtype, observed during 2000–2001, support this view. Nevertheless, it is not uncommon for new information about cases to emerge, often several years later, thus rendering the original epidemiological categorization incorrect. Accordingly, it is possible that some of the nine non-B subtype cases that reported a UK-only exposure acquired their infection outside the United Kingdom.

In contrast to the uncertainty surrounding the veracity of behavioural/demographic information, we are confident that the subtype analysis yielded accurate results. Validation of such accuracy was achieved through comparing the subtype status of both env and gag regions. In the 82% of instances where sequence of sufficient quality was generated in both regions, there was subtype concordance.

The surveillance approach now used in Scotland can be compared with that adopted in England. In the former, genotyping of HIV from newly diagnosed individuals is undertaken, while in the latter, an algorithm employing serotyping followed by molecular techniques of HIV antibody-positive specimens obtained through the unlinked, anonymous testing of genito-urinary medicine clinic attendees is performed. The advantage of the method used in England is its facility to monitor subtype distribution among both diagnosed and undiagnosed HIV-infected persons. In Scotland, monitoring is confined to infected persons who seek a diagnostic test. The unlinked, anonymous serotyping method offers the truer reflection of subtype epidemiology but that is offset by its limitations concerning the collection of demographic and behavioural information such as age band, gender, country of birth and sexual orientation. Because of the anonymity factor, it is impossible to trace a particular patient to check on, for example, his/her geographical exposure details. Subject to consent, however, it is possible to interview a patient if he/she has been subtyped in the context of a named HIV-1-positive diagnosis.

The serotyping approach is cheaper and allows for larger numbers of specimens to be tested. 'Serosubtype', however, does not always correspond to 'geno-subtype', perhaps in as many as 10% of cases [15, 16]. Whilst the use of genotyping is limited to specimens with detectable viral loads it is generally regarded as the gold standard and if carried out in the pol region, has the advantage that the anti-retroviral drug sensitivities of the virus can be inferred. It should be noted that neither approach is able to differentiate new infections from long-established ones. Indeed, no routine test to accurately identify incident HIV-1 infections is currently available.

In the first year of surveillance in Scotland, samples from approximately 90% of all diagnoses were obtained and, of these, 90% yielded sufficient virus to generate sequence in either or both regions. Where viral loads were known for samples that failed to amplify, they were invariably near or below the undetectable level. Amplification failures due to undetectable virus were more common in samples from MSM than other risk groups. The not uncommon occurrence of MSM having been previously diagnosed elsewhere (in particular England) and commenced on highly active anti-retroviral therapy before 're-diagnosis' on entry into clinical care in Scotland, is a likely explanation for this observation. For other samples that failed to amplify, the volume of starting materials supplied was below the optimum for extraction (results not shown). It was of concern that samples which failed to amplify harboured a subtype which could not be recognized by the primers used. The variety of subtypes found, however, did not support this hypothesis.

In a recently published study involving the serotyping of HIV antibody-positive specimens identified through the unlinked, anonymous HIV-1 testing of genito-urinary medicine-clinic attendees in England, it was concluded that, for 1997, 75% of all HIV-1

infections were caused by subtype B viruses [15]. This figure was not evenly spread amongst risk categories, with just over half of those infections acquired through the heterosexual route being non-B subtypes. While we also found that the subtype B viruses were responsible for the great majority (93%) of infections in MSM and IDUs, non-B subtypes were the cause of 74% of heterosexually acquired HIV-1 infections in Scotland. The corresponding non-B subtype rate for diagnoses made between 1995 and 1997 was 40% [17]. Subtype C still represented the dominant non-B subtype, a finding that reflects its predominance globally. Although most non-B infections originate in Africa, there is evidence that the dissemination of such viruses into the indigenous population may have occurred – a finding which contrasts with observations made by the same investigators in the mid-1990s [17] – and that such transmission may not be restricted to heterosexuals.

The principal reported mode of transmission for HIV-1 among persons diagnosed in the United Kingdom since 1999 has been heterosexual contact [18]. Should non-B viruses be inherently better adapted to spread through unprotected heterosexual intercourse than subtype B ones, their presence in the indigenous heterosexual community could herald a scenario whereby HIV-1 might no longer be considered a relatively rare infection among the general heterosexual population. Furthermore, should it transpire that different subtypes have different clinical outcomes or responses to anti-retroviral therapy, knowledge of subtype would directly influence patient management. Clearly, the spread of HIV-1 subtypes must continue to be monitored closely.

REFERENCES

1. Robertson DL, Anderson JP, Bradac JA, et al. HIV-1 nomenclature proposal. Science 2000; **288**: 55–56.
2. Sharp PM, Robertson DL, Gao F, Hahn BH. Origins and diversity of human immunodeficiency viruses. AIDS 1994; **8**: S27–S42.
3. UNAIDS (http://www.unaids.org). Accessed March 2004.
4. Expert Group of the Joint United Nations Programme on HIV/AIDS. Implications of HIV variability for transmission: scientific and policy issues. AIDS 1997; **11**: UNAIDS 1–15.
5. Mastro TD, Kunanusont C, Dondero TJ, Wasi C. Why do subtypes segregate among persons with different risk behaviours in South Africa and Thailand? AIDS 1997; **11**: 113–116.
6. Kaleebu P, French N, Mahe C, et al. The role of HIV-1 envelope subtypes A and D on disease progression in a large cohort of HIV-1 positive individuals in Uganda. J Infect Dis 2002; **185**: 1244–1250.
7. Peeters M, Vincent R, Peret J-L, et al. Evidence for differences in MT2 cell tropism according to genetic subtypes of HIV-1: syncytium-inducing variants seem rare among subtype C HIV-1 viruses. J Acquir Immune Defic Syndr 1999; **20**: 115–121.
8. Deschamp D, Apetrei C, Collin G, et al. Naturally occurring decreased susceptibility of HIV-1 subtype G to protease inhibitors. AIDS 1998; **12**: 1109–1111.
9. Alaeus A, Lidman K, Sonnerborg A, Albert J. Subtype-specific problems with quantification of plasma HIV-1 RNA. AIDS 1997; **11**: 859–865.
10. Yirrell DL, Kaleebu P, Morgan D, et al. Inter- and intra-genic intersubtype HIV-1 recombination in rural and semi-urban Uganda. AIDS 2002; **16**: 279–286.
11. Los Alamos National Laboratory, Los Alamos, New Mexico, 87545 USA. HIV databases (http://hiv-web.lanl.gov). Accessed March 2004.
12. Smith SW, Overbeek R, Woese CR, Gilbert W, Gillevet PM. The genetic data environment; an expandable GUI for multiple sequence analysis. Comput Applic Biosci 1994; **10**: 671–675.
13. Hall T. BioEdit Sequence Alignment Editor, version 4.5.8, Department of Microbiology, North Carolina State University, USA (http://www.mbio.ncsu.edu). Accessed March 2004.
14. Van de Peer Y, De Wachter R. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput Applic Biosci 1994; **10**: 569–570.
15. Parry JV, Murphy G, Barlow KL, et al. National surveillance of HIV-1 subtypes for England and Wales: design, methods and initial findings. J Acquir Immune Defic Syndr 2001; **26**: 381–388.
16. Kaleebu P, Yirrell D, French N, et al. An improved algorithm for detecting HIV type 1 subtypes in a primary laboratory in Uganda. AIDS Res Hum Retro 2000; **16**: 621–625.
17. Yirrell DL, Goldberg DJ, Whitelaw J, McSharry C, Raeside F, Codere G. Viral subtype and heterosexual acquisition of HIV in Scotland. Sex Trans Dis 1999; **75**: 392–395.
18. Public Health Laboratory Service (http://www.hpa.org.uk). Accessed March 2004.