

Consensus drug resistance mutations for epidemiological surveillance: basic principles and potential controversies

Robert W Shafer^{1*}, Soo-Yon Rhee¹ and Diane E Bennett²

¹Division of Infectious Diseases, Stanford University, Stanford, CA, USA

²World Health Organization, Geneva, Switzerland

*Corresponding author: E-mail: rshafer@stanford.edu

Programmes that monitor local, national and regional levels of transmitted HIV-1 drug resistance inform treatment guidelines and provide feedback on the success of HIV-1 treatment and prevention programmes. The World Health Organization (WHO) has established a global programme for genotypic surveillance of HIV-1 drug resistance and has recommended the adoption of a consensus definition of genotypic drug resistance. Such a definition is necessary to accurately compare transmitted drug resistance rates across geographical regions and time periods.

HIV-1 diversity and the large number of mutations associated with antiretroviral drug resistance complicate the development of a consensus definition for genotypic drug resistance. This paper reviews the data that must be considered to determine which of the many HIV-1 drug resistance mutations are likely to be both sensitive and specific indicators of transmitted drug resistance. The process used to create a previously published list of drug resistance mutations for HIV-1 surveillance is reviewed and alternative approaches to this process are discussed.

Introduction

The WHO has established a global programme for genotypic surveillance of transmitted HIV-1 drug resistance in representative untreated populations so that a rise in the prevalence of resistance can be detected at an early stage [1,2]. This programme is confronted by several logistical and operational challenges including the identification of representative populations for sampling and the development of the infrastructure for recruiting infected individuals, processing clinical samples and standardizing laboratory procedures. An additional challenge caused by HIV-1 genetic diversity and the many known HIV-1 drug resistance mutations is the development of a consensus definition of genotypic drug resistance. Such a definition is needed to accurately compare transmitted resistance rates across different regions and time periods [1–5].

We recently outlined considerations for identifying surveillance drug resistance mutations (SDRMs) and developed a consensus list of 80 SDRMs at 39 positions in HIV-1 reverse transcriptase (RT) and protease [5]. Mutations were included on the list if they met three criteria: they caused or contributed to HIV-1 drug resistance; they did not occur in the absence of

antiretroviral (ARV) drug pressure; and they were valid for all subtypes. The mutation list itself was designed to be simple, unambiguous and parsimonious. We refer to that list here as the provisional SDRM list. Figure 1 depicts the steps used to create this list. Although this approach appears reasonable, using it to develop an optimal consensus list is not straightforward and might in some cases be controversial. In this article, we provide data and arguments in support of seven propositions pertaining to the selection and use of drug resistance mutations for studies of transmitted HIV-1 drug resistance.

The first proposition provides the rationale for using genetic sequence data as the primary indicator of drug resistance. The next four propositions address the manner in which SDRMs were chosen for inclusion in the provisional SDRM list as well as how future mutations should be assessed when creating updates to the list. The sixth proposition introduces the concept that molecular phylogenetic factors might become increasingly important for choosing SDRMs and for analysing drug resistance surveillance data. The final proposition argues that an online program that can analyse a set of sequences using one or more

SDRM lists will standardize resistance prevalence calculations and facilitate meta-analyses of surveillance data collected by different investigators at different times.

Genotypic sequences are the optimal form of resistance data for surveillance studies

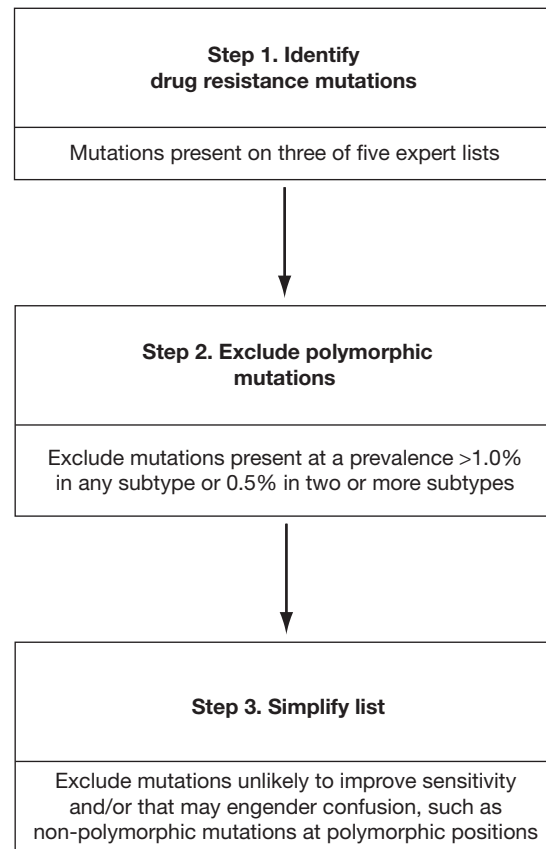
HIV-1 replication is highly error-prone. Within an infected individual, HIV-1 exists as a mixture of innumerable variants that diverge from the initially transmitted virus variant. Within human populations, high rates of polymorphism are found even in the relatively conserved molecular targets of therapy – RT and protease. HIV-1 diversity and the large number of mutations associated with ARV resistance add to the host of complications that must be addressed when developing a consensus genotypic definition for drug resistance.

However, despite these difficulties, genotypic resistance testing has three major advantages over phenotypic resistance testing for HIV-1 drug resistance surveillance. First, unlike bacterial or mycobacterial susceptibility testing, HIV-1 susceptibility testing is difficult to perform and even more difficult to perform well. Only a few laboratories are capable of performing HIV-1 susceptibility testing and only one laboratory has demonstrated a high level of reproducibility in tests of all three classes of RT and protease inhibitors [6,7]. Moreover, phenotypic susceptibility testing is done using proprietary methods. No reliable standardized test has been developed that would allow testing to be performed in the countries and regions in which surveillance is being performed.

In contrast, genotypic susceptibility testing has become highly reproducible. A few early studies performed in the 1990s reported inconsistent results among different laboratories performing genotypic susceptibility testing [8,9]. However, the inconsistencies in these early studies were attributed to a small set of the participating laboratories [8]. Nearly all subsequent studies have described high levels of reproducibility among laboratories performing HIV-1 genotypic resistance testing [10–14].

The second advantage of genotypic testing is its greater sensitivity at detecting drug resistance in previously untreated persons. A significant proportion of persons with transmitted resistance have revertant mutations (for example, T215C/D/E/I/S/V) that indicate transmitted resistance but do not reduce drug susceptibility [15,16]. Other persons with transmitted resistance have drug resistance mutations co-circulating with wild-type virus variants that emerge as a result of back mutation during the months to years following initial infection [17,18]. Genotypic resistance

Figure 1. An overview of the steps used to create the provisional list of surveillance drug resistance mutations



Beneath each step is a description of the specific approach used. Possible alternatives to the approaches used to derive the provisional surveillance drug resistance mutations (SDRMs) list are reviewed in this article.

testing detects resistance in these individuals. In contrast, phenotypic resistance testing will detect a change in susceptibility only if the mutant variant is present at sufficiently high levels relative to wild-type. The deleterious effect of mutations on specific drugs can also be masked by the presence of combinations of mutations with antagonistic effects on susceptibility. Thus, the effect of thymidine analogue mutations (TAMs) may not be appreciated phenotypically in the presence of the mutation M184V, which increases zidovudine, stavudine and tenofovir susceptibility [19]. Likewise, the effect of non-nucleoside reverse transcriptase inhibitor (NNRTI) resistance mutations might be minimized in the presence of multiple TAMs, which increase NNRTI susceptibility [20,21].

Finally, genotypic resistance test results provide molecular epidemiological data in addition to drug resistance data. Such data are useful not only for

detecting technical problems such as laboratory contamination and sample mix-up but, more importantly, for gaining insight into the population genetics of transmitted resistance.

Identifying drug resistance mutations requires expert judgment

The literature on HIV drug resistance mutations is vast. Many mutations in protease and RT have been linked to drug resistance by one of three criteria: selection *in vitro* or *in vivo* by an ARV drug; association with decreased *in vitro* susceptibility; or association with a decreased virological response to therapy [22–24]. The strength of the supporting evidence for different drug resistance mutations is highly variable. Some mutations fulfill all three criteria solidly; others fulfill only one or two criteria. The association of other mutations with ARV treatment, reduced *in vitro* susceptibility or decreased response to ARV therapy is supported only by the flimsiest of statistical evidence.

We adopted an ad hoc approach for constructing the provisional SDRM list that defined drug resistance mutations as those present in three or more of five published lists of drug resistance mutations [23,25–28]. Although this simplified the process for choosing among a large number of potential drug resistance mutations, this approach is likely to be suboptimal for future updates. One reason is that none of the published lists was designed for drug resistance surveillance. Two were designed to be primarily educational [23,27] and three were designed to assist genotypic drug resistance test interpretation [25,26,28]. Indeed, it was necessary to make exceptions in the provisional list for two mutations that were not present in three or more lists: the RT codon 215 revertant mutations (for example, T215D/C/E/S/I/V) and the protease mutation V82M. Codon 215 revertant mutations, which occur in approximately 3% of newly diagnosed persons in the US, are one of the strongest molecular indicators of HIV resistance transmission [15,16]. V82M, although uncommon in protease of subtype B isolates, is a common cause of protease inhibitor (PI) resistance in subtype G isolates [29].

In conclusion, we do not believe it is possible to use a single systematic or automated approach to identify all HIV-1 mutations that cause or contribute to resistance. Many drug resistance mutations do not meet all three criteria (selection by therapy, reduced *in vitro* susceptibility and reduced virological response to therapy), and the supporting evidence for each drug resistance mutation is varied. For these reasons, expert judgment such as that provided by an expert panel, formed specifically to develop a list of surveillance mutations, will remain necessary to weigh the lines of

evidence suggesting that a mutation causes or contributes to drug resistance.

Low-level polymorphism may reflect transmitted resistance or naturally occurring polymorphism

Some drug resistance mutations occur commonly in the absence of selective drug pressure. Including such polymorphic mutations in a list of SDRMs would clearly overestimate the levels of transmitted resistance. Among the most well known mutations of this type are the accessory PI resistance mutations at positions 10, 20, 36, 63, 71, 77 and 93, which – depending on the subtype – occur in 5% to >50% of untreated persons. Furthermore, the accessory nucleoside reverse transcriptase (NRTI) mutations E44D and V118I and the minor NNRTI resistance mutations A98G, V108I and V179D occur in between 1% and 2% of viruses from untreated persons.

Some non-polymorphic drug resistance mutations, however, may appear polymorphic in analyses that contain a large proportion of untreated persons who are primarily infected with resistant viruses – a frequent occurrence in areas where ARV drugs have been available for many years. These mutations are important indicators of transmitted resistance and should not be excluded from a list of SDRMs. When creating the provisional SDRM list, we used a three-step approach to minimize the chance that transmitted resistance obscured truly non-polymorphic drug resistance mutations.

First, we compiled nearly all publicly available HIV-1 RT and protease sequences in GenBank obtained from persons documented to be ARV-naïve in the publications associated with the GenBank submissions. This compilation included RT sequences from 9,417 persons and protease sequences from 11,791 ARV-naïve persons. Non-subtype B sequences made up 52% of the RT sequences and 49% of the protease sequences. Overall, 29% of these sequences were from Africa, 20% from North America, 19% from South America, 16% from Europe, and 16% from Asia. A list of the studies examined for this analysis can be found at <http://hivdb.stanford.edu/cgi-bin/reference.cgi>.

Next, we excluded sequences likely to be at high risk for transmitted resistance: those from persons who were infected in regions where transmitted resistance is common, and those from untreated persons that contained two or more drug resistance mutations at known non-polymorphic positions. Finally, we calculated the prevalence of each drug resistance mutation in each of the eight most common subtypes. Those mutations occurring at a prevalence of $\leq 0.5\%$ in all subtypes were considered non-polymorphic. However, because insufficient data were available for some subtypes,

mutations that were present at a prevalence >0.5% but <1.0% in no more than one non-B subtype were provisionally considered an SDRM.

Tables 1 to 3 show the prevalence of each of the SDRMs among untreated and treated persons in each of the eight most common subtypes. Despite the addition of new studies published during the past year, all but one of the mutations on the provisional SDRM list has remained nonpolymorphic according to the current working definition. M41L, one of the most commonly transmitted drug resistance mutations, occurred in 0.9% of ARV-naive subtype D (2/224) and subtype G (2/209) isolates. Five other mutations occurred with a prevalence between 0.6% and 0.9% in a single non-B subtype.

About one-half of the SDRMs on the provisional list occur at a prevalence of 0.1% to 0.5% in one or more subtypes (Tables 1–3). For example, the NRTI resistance mutation M41L, the NNRTI resistance mutation K103N, and the PI resistance mutation L90M occur in 0.4%, 0.4% and 0.2% of untreated persons with subtype B viruses, respectively. The presence of such mutations in untreated persons probably reflects a combination of transmitted resistance, present despite the exclusions described above, and naturally occurring low levels of polymorphism. For example, the NRTI resistance mutation K70R, which has been reported in 0.2% of untreated persons infected with subtype B and CRF01_AE viruses, had been reported in isolates even before the use of ARV drugs [30,31].

Table 1. HIV-1 NRTI surveillance drug resistance mutations: prevalence in RTI-naive and RTI-treated persons according to subtype

Pos [†]	AA [§]	RTI-naive persons [*]								RTI-treated persons [†]								
		A	AE	AG	B	C	D	F	G	A	AE	AG	B	C	D	F	G	
		519	1,043	720	4,545	1,950	224	207	209	296	746	362	10,942	793	261	289	630	
M41	L	0	0	0.3	0.4	0.1	0.9 [‡]	0.5	0.9 [‡]	8.3	32	18	37	14	16	31	28	
K65	R	0	0	0	0	0.1	0	0	0	2.7	3.2	2.8	1.8	4.4	0.8	1.4	4.4	
D67	G	0	0	0	0	0.2	0	0	0	1.4	1.1	1.1	1.5	0.5	0.8	0.7	1.3	
	N	0	0	0	0	0	0	0	0	16	44	22	28	22	19	26	24	
	Del	0	0	0	0	0	0	0	0	0	0	0	0.1	0	0	0	0	
T69	D	0.2	0	0	0	0	0	0	0	1.4	5.4	2.8	6.8	2.6	1.6	5.2	1	
	Ins	0	0	0	0	0	0	0	0	0	0.5	0	1.2	0	0.8	0	0	
K70	R	0	0.2	0	0.2	0	0	0	0	19	30	17	22	16	14	24	15	
L74	V	0	0	0	0	0.1	0	0	0.5	1.4	4.6	2.8	5.8	3.2	5.4	4.5	6.2	
V75	A	0	0	0	0	0.1	0	0	0	0	0.8	0.6	0.2	0.3	0	0	0.3	
	M	0	0.3	0	0	0	0	0	0	1.4	3.8	0	1.9	0.3	0.8	2.1	1.8	
	S	0	0	0	0	0	0	0	0	0	0	0	0.1	0	0	1.4	0	
	T	0	0	0	0	0	0	0	0	0.7	2.4	2.2	0.9	0.8	0.8	0	0.6	
F77	L	0	0.1	0	0	0	0	0	0	0.7	2.1	1.1	1.8	1.3	0	0.7	0.6	
Y115	F	0	0	0	0	0	0	0	0	0.7	1.1	1.1	1.2	1.5	0	0	0.3	
F116	Y	0	0	0	0	0	0	0	0	0.7	3.2	1.7	2	2.5	0.8	0.7	1.6	
Q151	M	0	0	0	0	0	0	0	0	0.7	5.1	2.8	2.5	2.9	0.8	0.7	2.5	
M184	I	0	0	0	0	0	0.5	0	0	0	2.7	1.2	1	0.3	0	0	0.7	
	V	0	0.2	0.1	0.2	0.1	0.5	0	0	44	28	44	45	39	38	38	58	
L210	W	0	0	0	0	0	0	0	0	3.9	20	6.4	23	5.8	7.5	11	4.8	
T215	C	0	0	0	0.2	0.1	0	0	0	0	0	0.6	0.6	0	2	1.8	0	
	D	0	0	0	0.3	0.1	0	0	0	0	0	0	0.7	0	0	0.7	0.3	
	E	0	0	0	0.1	0	0	0	0	0	0	0	0.1	0	2	0	0.3	
	F	0	0	0	0	0	0	0	0	11	21	8	8.7	3.9	5.1	14	6.1	
	I	0.4	0	0	0	0	0	0	0	5	0.5	1.1	0.6	0.3	0	1.1	0.3	
	S	0	0	0	0.3	0.1	0	0	0	0	0.3	1.1	0.5	0	0	0	0	
	V	0	0	0	0	0	0	0	0	1.2	0.3	0	0.7	0.3	0	0.7	0.3	
	Y	0.2	0	0.1	0	0.1	0	0	0	8.3	25	16	36	19	14	27	26	
	K219	E	0	0	0	0	0	0	0	0	7.1	4.6	8	3.3	4.9	1	6.7	7.1
		Q	0.2	0.1	0.1	0	0	0.9 [‡]	0	0	12	25	7.1	14	5	9	11	6.8
R		0	0.1	0	0.1	0.1	0	0	0	0	0.8	0	1.7	0.5	1	1.4	1.3	

*The header for columns 3–10 contains the number of sequences in the HIV Drug Resistance Database from untreated persons by subtype. [†]The header for columns 11–18 contains the number of sequences from treated persons according to subtype. The row entries contain the mutation prevalence (%) according to subtype and treatment using a set of publicly available sequences updated on the 1 October 2007. [‡]Pos, reverse transcriptase position preceded by the consensus residue. [§]AA, surveillance drug resistance mutation (Del, deletion; Ins, insertion). [¶]Mutations with a prevalence >0.5%. RTI, reverse transcriptase inhibitor; SDRM, surveillance drug resistance mutation.

In conclusion, large numbers of sequences from untreated persons infected with viruses of different subtypes are required to distinguish polymorphic from non-polymorphic drug resistance mutations. There is no definitive approach for distinguishing low levels of polymorphism arising from natural occurrence versus transmitted resistance, because even in recently infected persons the genotype and treatment history of the source of infection are rarely known. Nonetheless, we believe that the heuristic approaches described here, which we used to develop the provisional SDRM list, represent a logical way to determine which drug resistance mutations are non-polymorphic.

Creating a single SDRM list applicable to all subtypes is both necessary and feasible

Most data on the genetic mechanisms of HIV-1 drug resistance arose from studies of subtype B viruses, the predominant subtype in North American and Europe. During the past few years, however, the literature and the number of published sequences from untreated and treated persons with non-subtype B viruses have expanded. Indeed, the number of untreated individuals with non-subtype B sequences for the seven most common non-B subtypes has increased in the published literature by 17% for RT (4,265–4,998 individuals) and

14% for protease (5,097–5,807 individuals) since the provisional SDRM list was created (Tables 1–3).

Analyses of this sequence data reveals that although polymorphism rates differ between subtypes, positions that are polymorphic in one subtype are generally polymorphic in all subtypes. Conversely, positions that are non-polymorphic in one subtype are generally non-polymorphic in all subtypes. This similarity of wild-type sequences from different subtypes is consistent with the large number of studies showing that currently available RT and protease inhibitors are equally active against all HIV-1 group M subtypes. Indeed, there are no published treatment guidelines that recommend different treatments for viruses belonging to different subtypes.

There are a few examples, however, of PI resistance mutations that – in the dataset currently available for analysis – are polymorphic in some subtypes but not in others. V11I, a mutation associated with darunavir resistance [32], occurs in 1.8% and 1.3% of untreated CRF02_AG and subtype G isolates, respectively, but in fewer than 1.0% of untreated isolates from the other six major subtypes. K20I occurs in >90% of untreated CRF02_AG and subtype G isolates, but is rarely found in untreated subtype B and C isolates. L33F occurs in ~1% of untreated subtype A and CRF01_AE isolates, but is rarely found in untreated

Table 2. HIV-1 NNRTI surveillance drug resistance mutations: prevalence in RTI-naïve and RTI-treated persons according to subtype

Pos [†]	AA [§]	RTI-naïve persons*									RTI-treated persons [†]						
		A	AE	AG	B	C	D	F	G	A	AE	AG	B	C	D	F	G
		519	1043	720	4,545	1,950	224	207	209	296	746	362	10,942	793	261	289	630
L100	I	0	0	0	0	0	0	0	0	2	1.9	1.7	1.6	1	1.5	2.4	2.6
K101	E	0	0	0	0.1	0.1	0	0	0	3.7	5.7	6.7	2	3.8	2.3	4.2	1.6
N103	N	0.2	0	0	0.4	0.2	0	0	0	18	13	26	17	18	17	24	37
	S	0	0	0	0	0.1	0	0	0	0	0	0	0.3	0.3	0	0	0.3
V106	A	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0	0.7	1.3
	M	0	0	0	0	0	0	0	0	0	1.1	0.6	0.2	9	0.8	1.4	0.3
Y181	C	0.2	0.1	0	0.1	0.1	0	0	0	8.3	9.4	7.8	7.9	9	5.4	15	12
	I	0	0.1	0	0	0	0	0	0	0.7	1.1	0	0.4	0.5	0	0.7	0.3
Y188	C	0	0	0.1	0	0	0	0	0	0	0	0	0.1	1	0.8	0	0.3
	H	0	0	0	0	0.1	0.5	0	0	0.7	0	0	0.2	0	0	0	0
G190	L	0	0.2	0	0	0.1	0	0	0	1.4	4.6	1.7	1.6	3.3	0.8	2.1	4.2
	A	0	0.1	0	0	0.1	0.5	0	0	11	13	4.5	4.5	9.2	7.7	4.6	5.3
	E	0	0	0	0	0.1	0	0	0	0	0.5	0	0.3	0.3	0	0.7	0.3
	Q	0	0	0	0	0	0	0	0	0	0	0	0.1	0	0	0	0
P225	S	0	0	0	0	0	0	0	0	0.7	1.3	5.6	1.1	1.3	0.8	1.4	0.3
	H	0	0	0	0	0	0	0	0	0	1.4	3.5	0.9	0.8	1.4	2.2	4
M230	L	0	0	0	0	0	0	0	0.8	0.3	1.8	0.5	1.3	0	0.8	1	
P236	L	0	0	0.3	0.1	0	0	0	0	0	0	0	0.2	0.6	0	0	0

*The header for columns 3–10 contains the number of sequences in the HIV Drug Resistance Database from untreated persons by subtype. [†]The header for columns 11–18 contains the number of sequences from treated persons according to subtype. The row entries contain the mutation prevalence (%) according to subtype and treatment using a set of publicly available sequences updated on the 1 October 2007. [†]Pos, reverse transcriptase position preceded by the consensus residue. [§]AA, surveillance drug resistance mutation. RTI, reverse transcriptase inhibitor; SDRM, surveillance drug resistance mutation.

Table 3. HIV-1 PI surveillance drug resistance mutations: prevalence in PI-naive and PI-treated persons according to subtype

Pos [‡]	AA [§]	PI-naive persons*								PI-treated persons [†]							
		A	AE	AG	B	C	D	F	G	A	AE	AG	B	C	D	F	G
		907	923	1,161	6,059	1,684	400	330	327	97	94	153	6,590	264	145	224	185
L24	I	0	0	0.1	0	0.1	0	0	0	5	0	1.3	7.5	1.5	8.5	11	2.2
D30	N	0	0	0	0	0	0	0	0	0	0	0	6.9	8.8	7.7	1.4	0.5
V32	I	0	0	0.1	0	0	0	0	0	0	0	0.7	7.3	0.4	8.3	2.2	0
M46	I	0.2	0.6 [¶]	0.1	0.2	0.1	0	0.3	0	19	4.5	17	29	12	26	13	17
I47	A	0	0	0	0	0	0	0	0	0	0	0	0.5	0.4	0.7	0.4	0
	V	0	0	0	0	0.1	0.2	0	0.6 [¶]	2	0	3.9	6.5	0.4	4.9	3.6	1.6
G48	V	0	0	0	0	0	0	0	0	3	8.6	1.3	5.7	1.5	4.1	4.1	2.7
I50	L	0	0	0	0	0	0	0	0	0	0	0	0.4	1.9	0	0	0
	V	0	0	0	0	0.1	0	0	0	0	0	0	2.5	0.4	3.4	2.7	1.6
F53	L	0	0.1	0	0	0.1	0	0	0.3	5	3.2	5.3	6	0.4	5.5	8.6	8.7
I54	A	0	0	0	0	0	0	0	0	1	1.1	0	1.6	1.1	2.2	1.4	1.7
	L	0	0	0	0	0	0	0	0	1	1.1	0	4	0.4	2.2	0	0
	M	0	0	0	0	0	0	0	0	0	0	0	4.2	0	3.6	0.9	0.6
	S	0	0	0	0	0	0	0	0	2	0	0.7	1.3	0	0.7	0.5	0
	T	0.1	0	0	0	0.1	0	0	0	2	0	0	1.4	0	1.4	1.4	0
	V	0	0	0	0	0	0	0	0	19	8.8	19	31	9.6	34	42	39
G73	A	0	0	0	0	0	0	0	0	0	0	0	0.6	0	0.7	0	0.5
	C	0	0	0	0	0	0	0.3	0	0	0	0	1.8	0	0	0.5	0
	S	0	0	0.1	0	0.1	0	0.3	0	3	2.2	1.3	11	2.3	9.1	1.8	0
	T	0	0	0	0	0	0	0	0	0	0	0	4	0	2.8	0.5	0
V82	A	0	0	0	0	0.1	0	0	0	11	11	8.1	29	12	28	41	3.3
	F	0	0	0	0	0	0	0	0	6	3.3	1.3	2.1	0.4	0.7	1.8	1.7
	M	0	0	0	0	0	0	0	0	0	0	0	0.2	0	0.7	0	3.3
	S	0	0	0	0	0	0	0	0	0	0	1.3	1.5	0.8	2.1	0.9	3.3
	T	0	0	0	0	0	0	0	0	1	1.1	1.3	3.4	1.2	4.3	2.7	13
I84	A	0	0	0	0	0	0	0	0	0	0	0.7	0.2	0	0	0	0
	C	0	0	0	0	0	0	0	0	0	0	0.7	0.2	0.4	0	0	0
	V	0	0	0	0	0	0	0	0	10	7.8	14	21	4.9	15	2.7	6.6
N88	D	0	0	0	0	0	0	0	0	1	0	0	5.3	5.4	5.6	2.2	0.5
	S	0	0	0.1	0	0.1	0	0	0.6 [¶]	2	3.2	1.3	1.1	3.1	0	3.1	1.1
L90	M	0.1	0	0.1	0.2	0.1	0	0.3	0.6 [¶]	12	9.8	19	44	24	31	21	39

*The header for columns 3–10 contains the number of sequences in the HIV Drug Resistance Database from untreated persons by subtype. [†]The header for columns 11–18 contains the number of sequences from treated persons according to subtype. The row entries contain the mutation prevalence (%) according to subtype and treatment using a set of publicly available sequences updated 1 October 2007. [‡]Pos, protease position preceded by consensus residue. [§]AA, surveillance drug resistance mutation. [¶]Mutations with a prevalence >0.5%. PI, protease inhibitor; SDRM, surveillance drug resistance mutation.

isolates belonging to other subtypes. M46L occurs in 1.3% of subtype G isolates, but not in other subtypes. T74S, a mutation recently shown to contribute to nelfinavir resistance [33], occurs in 8.1% of untreated subtype C isolates but is rarely found in untreated subtype B isolates.

The genetic mechanisms of ARV resistance are also highly similar among different subtypes (that is, the mutations that cause resistance in subtype B viruses also cause drug resistance in non-B viruses) [34]. There are two notable exceptions. First, although the NNRTI resistance mutation V106M is one of the most common mutations emerging in subtype C viruses during NNRTI therapy [35,36], it occurs rarely in subtype B viruses subjected to the same drug pressure. Second, the PI resistance mutation V82M

occurs commonly in subtype G viruses exposed to PIs [29], but rarely in subtype B viruses under the same conditions. Differences in the proportions of certain mutations occurring during PI therapy, such as the frequency of L90M relative to D30N in persons receiving nelfinavir, have also been reported [33,37–40]. Such differences, however, do not influence the composition of the SDRM list, because both D30N and L90M are included in the list.

These observations support the concept that a single SDRM list applicable to all subtypes can be created. The creation of different SDRM lists for different subtypes is unnecessary and would complicate surveillance efforts, because in many regions multiple subtypes are found. Finally, an increasing number of new infections appear to be caused by intersubtype recombinants, and determining

the subtype of HIV-1 RT and protease in these recombinants can be difficult [41].

An SDRM list should be refined by examining sequences from recently infected persons

An SDRM list does not have to contain all non-polymorphic drug resistance mutations. Rather, the list should be the shortest possible list that reliably and consistently assesses the prevalence of drug resistance across different populations. Each additional mutation included on an SDRM list poses a slight risk of inflating the background prevalence of resistance because, as noted above, even highly conserved mutations can occur naturally in 0.1% to 0.5% of sequences. Expert judgment such as that provided by an expert panel could be used to exclude mutations that are unlikely to influence prevalence calculations either because the mutations are rare or because they occur only in combination with other drug resistance mutations.

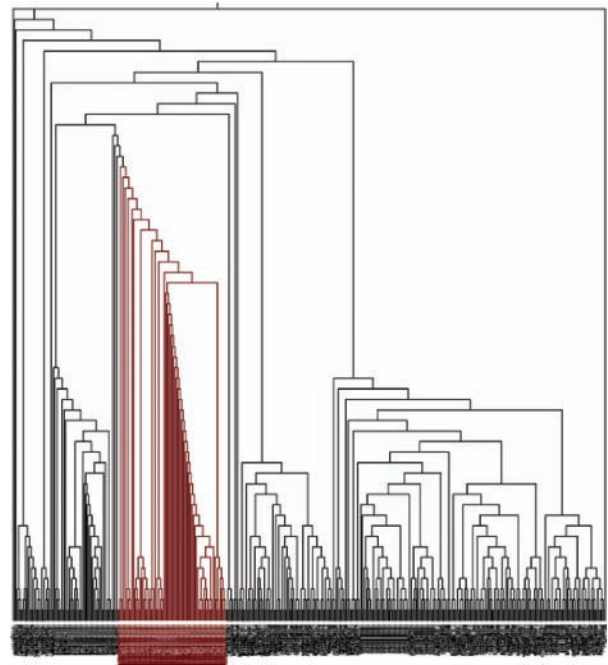
The vast majority of published sequences are either from ARV-experienced persons or from ARV-naive persons in whom the source and time of infection is not known. Sequences from recently infected persons or from persons with a known source of infection have infrequently been published. However, such sequences, as well as the many sequences that will be obtained during surveillance, will make it possible to fine-tune an SDRM list. In particular, extraneous, uninformative mutations that do not occur alone in recently infected individuals can be removed. Conversely, rare mutations not currently on the SDRM list that are found to be specific indicators of resistance transmission can be added.

Clusters of phylogenetically linked variants could bias SDRM prevalence rates

Epidemiologists in Sweden recently described a transmission cluster of six newly diagnosed individuals who presented between 1993 and 1996 with the NRTI resistance mutation M41L [42]. This mutation presumably arose in one person receiving therapy and was then transmitted to several untreated persons. Likewise, the otherwise non-polymorphic NRTI resistance mutation A62V is found in >100 of ~500 subtype A isolates from untreated individuals. However, all of the subtype A variants with A62V are from countries of the former Soviet Union and are closely related phylogenetically – differing from one another by a mean of 1.0% \pm 0.7% of their nucleotides (Figure 2) [43].

Similar to most SDRMs, stably transmitted resistant variants may influence the success of initial ARV therapy and are, therefore, of public health importance. However, in contrast to most SDRMs, stably transmitted

Figure 2. Neighbour-joining tree containing reverse transcriptase sequences from >500 untreated persons with viruses belonging to subtype A



The tree was created in PAUP 4.0 using the HKY85 substitution model with rate variation conforming to a gamma distribution. The branches highlighted lead to ~100 sequences with the reverse transcriptase (RT) mutation A62V. The cladogram (branches not drawn to scale) indicates that the clustering of viruses with this mutation results from a founder effect [43].

variants do not indicate the extent to which transmitted resistance results from ARV treatment failure. For example, the public health consequences of a drug resistance mutation occurring in five out of 100 persons may be different if the five resistant viruses are closely related phylogenetically – suggesting that they resulted from one ARV treatment failure – than if they resulted from five unrelated treatment failures.

Clusters of stably transmitted resistant variants are likely to increase as drug treatment and drug resistance surveillance programs expand. In particular, clusters are likely to be more apparent when the populations undergoing surveillance are geographically concentrated. Although such clusters may complicate SDRM updates by biasing the prevalence of drug resistance mutations in different populations, the possible public health significance of drug resistance clusters could be even more important. However, performing phylogenetic analyses of sequences obtained during surveillance and using these analyses to inform public health policy represents uncharted territory. Further discussion of this topic would be speculative and beyond the scope of this article.

An online program can standardize resistance prevalence calculations and facilitate meta-analyses across multiple studies

Comparisons of resistance prevalence results from previously performed studies are complicated by the fact that different studies often use different SDRM lists, as well as different methods for sampling untreated persons. Whereas not much can be done to standardize resistance prevalence results for differences in study design, it is possible to standardize resistance prevalence results for the use of different SDRM lists. Because sequences from different drug resistance surveillance studies are rarely made publicly available, the only way to compare resistance prevalence among populations is to use a computer program that allows epidemiologists to use the same SDRM list to analyse different datasets. Such a program can also ensure a consistent approach to handling missing data, such as incomplete or poor quality sequences.

The Calibrated Population Resistance (CPR) tool is an online program available on the HIV Drug Resistance Database website (<http://hivdb.stanford.edu>). It uses the provisional SDRM list to compute the prevalence of NRTI, NNRTI, PI, dual-class, and triple-class resistance in a set of RT and/or protease sequences from untreated persons submitted by the investigator. It also provides the prevalence of each SDRM in the set and an assessment of sequence quality. Most importantly, however, CPR allows users to select from several SDRM lists. For example, although the provisional SDRM list is the default list, users may select the commonly used list of major International AIDS Society USA (IAS-USA) drug resistance mutations. Moreover, when the SDRM list is updated, the original provisional list will remain an option so that all analyses are backward and forward compatible.

Conclusions

There are many challenges to performing epidemiologically sound studies of transmitted HIV-1 drug resistance. These challenges include identifying the most appropriate HIV-1 subjects, developing the local infrastructure for genotypic resistance testing, and determining how to optimally use surveillance results in the context of HIV-1 drug treatment and prevention programmes. However, researchers intimately involved either in measuring HIV-1 drug resistance within populations or in performing meta-analyses of published studies have realized that genotypically defining HIV-1 drug resistance is far from trivial. Indeed, for global analyses, the WHO will use a consensus definition of transmitted resistance based on a standard list; the same list will be recommended for use by countries in

their own programmes for surveillance of transmitted resistance [1,2].

In a previous manuscript, we described a step-by-step procedure for creating a list of SDRMs to serve as a genotypic definition of drug resistance [5]. This article discusses the merits of each criteria used to identify SDRMs and describes the types of decisions that are ultimately necessary for applying these criteria to publicly available HIV-1 sequences and the vast HIV-1 drug resistance literature. This article also introduces two new concepts that will become increasingly relevant as more drug surveillance studies are performed. First, HIV-1 RT and protease sequences often contain phylogenetically meaningful data that might influence how drug resistance surveillance data should be interpreted, although as noted in the text the use of molecular phylogeny in the setting of HIV-1 drug resistance surveillance has no precedent. Second, SDRM lists will inevitably be updated as new data and drugs become available. These updates will be most effective when they occur in the context of a publicly available online program that allows researchers possessing different datasets to perform identical analyses with either the original provisional SDRM list, future upgrades of this list or other published lists from previous studies.

Acknowledgements

RWS and SYR were supported in part by a grant from the Division of AIDS at the NIAID, NIH (AI068581). International Medical Press is acknowledged for their support with publishing this manuscript. Some of the authors are staff members of the WHO. The authors alone are responsible for the views expressed in this publication and they do not necessarily represent the decisions or stated policy of the WHO.

Disclosure statement

The authors declare that they have no conflicts of interest.

References

1. Bennett DE. The requirement for surveillance of HIV drug resistance within antiretroviral rollout in the developing world. *Curr Opin Infect Dis* 2006; **19**:607–614.
2. Bennett DE, Myatt M, Bertagnolio S, Sutherland D, Gilks CF. Recommendations for surveillance of transmitted HIV drug resistance in countries scaling up antiretroviral treatment. *Antivir Ther* 2008; **13** Suppl 2:25–36.
3. Pillay D. Current patterns in the epidemiology of primary HIV drug resistance in North America and Europe. *Antivir Ther* 2004; **9**:695–702.
4. Geretti AM. Epidemiology of antiretroviral drug resistance in drug naive persons. *Curr Opin Infect Dis* 2007; **20**:22–32.
5. Shafer RW, Rhee SY, Pillay D, *et al.* HIV-1 protease and reverse transcriptase mutations for drug resistance surveillance. *AIDS* 2007; **21**:215–223.

6. Parkin NT, Hellmann NS, Whitcomb JM, Kiss L, Chappey C, Petropoulos CJ. Natural variation of drug susceptibility in wild-type HIV-1. *Antimicrob Agents Chemother* 2004; **48**:437–443.
7. Zhang J, Rhee SY, Taylor J, Shafer RW. Comparison of the precision and sensitivity of the Antivirogram and PhenoSense HIV drug susceptibility assays. *J Acquir Immune Defic Syndr* 2005; **38**:439–444.
8. Schuurman R, Demeter L, Reichelderfer P, Tijnagel J, de Groot T, Boucher C. Worldwide evaluation of DNA sequencing approaches for identification of drug resistance mutations in the human immunodeficiency virus type 1 reverse transcriptase. *J Clin Microbiol* 1999; **37**:2291–2296.
9. Schuurman R, Brambilla D, de Groot T, *et al.* Underestimation of HIV type 1 drug resistance mutations: results from the ENVA-2 genotyping proficiency program. *AIDS Res Hum Retroviruses* 2002; **18**:243–248.
10. Shafer RW, Hertogs K, Zolopa AR, *et al.* High degree of interlaboratory reproducibility of human immunodeficiency virus type 1 protease and reverse transcriptase sequencing of plasma samples from heavily treated patients. *J Clin Microbiol* 2001; **39**:1522–1529.
11. Korn K, Reil H, Walter H, Schmidt B. Quality control trial for human immunodeficiency virus type 1 drug resistance testing using clinical samples reveals problems with detecting minority species and interpretation of test results. *J Clin Microbiol* 2003; **41**:3559–3565.
12. Grant RM, Kuritzkes DR, Johnson VA, *et al.* Accuracy of the TRUGENE HIV-1 genotyping kit. *J Clin Microbiol* 2003; **41**:1586–1593.
13. Eshleman SH, Hackett J, Jr., Swanson P, *et al.* Performance of the Celera Diagnostics ViroSeq HIV-1 Genotyping System for sequence-based analysis of diverse human immunodeficiency virus type 1 strains. *J Clin Microbiol* 2004; **42**:2711–2717.
14. Descamps D, Delaunay C, Masquelier B, *et al.* Repeated HIV-1 resistance genotyping external quality assessments improve virology laboratory performance. *J Med Virol* 2006; **78**:153–160.
15. Yerly S, Rakik A, De Loes SK, *et al.* Switch to unusual amino acids at codon 215 of the human immunodeficiency virus type 1 reverse transcriptase gene in seroconvertors infected with zidovudine-resistant variants. *J Virol* 1998; **72**:3520–3523.
16. Garcia-Lerma JG, Nidtha S, Blumoff K, Weinstock H, Heneine W. Increased ability for selection of zidovudine resistance in a distinct class of wild-type HIV-1 from drug naive persons. *Proc Natl Acad Sci U S A* 2001; **98**:13907–13912.
17. Little S, Dawson K, Hellmann N, DD Richman, Frost S. Persistence of transmitted drug-resistant virus among subjects with primary HIV infection deferring antiretroviral therapy. *Antivir Ther* 2003; **8**:S129.
18. Gandhi RT, Wurcel A, Rosenberg ES, *et al.* Progressive reversion of human immunodeficiency virus type 1 resistance mutations *in vivo* after transmission of a multiply drug resistant virus. *Clin Infect Dis* 2003; **37**:1693–1698.
19. Whitcomb JM, Parkin NT, Chappey C, Hellmann NS, Petropoulos CJ. Broad nucleoside reverse-transcriptase inhibitor cross-resistance in human immunodeficiency virus type 1 clinical isolates. *J Infect Dis* 2003; **188**:992–1000.
20. Whitcomb JM, Huang W, Limoli K, *et al.* Hypersusceptibility to non-nucleoside reverse transcriptase inhibitors in HIV-1: clinical, phenotypic and genotypic correlates. *AIDS* 2002; **16**:F41–47.
21. Shulman NS, Bosch RJ, Mellors JW, Albrecht MA, Katzenstein DA. Genetic correlates of efavirenz hypersusceptibility. *AIDS* 2004; **18**:1781–1785.
22. Rhee SY, Fessel WJ, Zolopa AR, *et al.* HIV-1 protease and reverse-transcriptase mutations: correlations with antiretroviral therapy in subtype B isolates and implications for drug resistance surveillance. *J Infect Dis* 2005; **192**:456–465.
23. Clark S, Calef C, Mellors JW. Mutations in retroviral genes associated with drug resistance. In *HIV Sequence Compendium 2006/2007*. Edited by: Leitner T, Foley B, Hahn B, *et al.* Los Alamos, NM, USA: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory; 2007. LA-UR 07-4826. pp.58–158.
24. Perno CF, Svicher V, Ceccherini-Silberstein F. Novel drug resistance mutations in HIV: recognition and clinical relevance. *AIDS Rev* 2006; **8**:179–190.
25. Van Laethem K, De Luca A, Antinori A, Cingolani A, Perna CF, Vandamme AM. A genotypic drug resistance interpretation algorithm that significantly predicts therapy response in HIV-1-infected patients. *Antivir Ther* 2002; **7**:123–129.
26. Rhee SY, Gonzales MJ, Kantor R, Betts BJ, Ravela J, Shafer RW. Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucleic Acids Res* 2003; **31**:298–303.
27. Johnson VA, Brun-Vezinet F, Clotet B, *et al.* Update of the Drug Resistance Mutations in HIV-1: 2005. *Top HIV Med* 2005; **13**:51–57.
28. Descamps D, Chaix ML, Andre P, *et al.* French national sentinel survey of antiretroviral drug resistance in patients with HIV-1 primary infection and in antiretroviral-naive chronically infected patients in 2001–2002. *J Acquir Immune Defic Syndr* 2005; **38**:545–552.
29. Camacho R, Godinho A, Gomes P, *et al.* Different substitutions under drug pressure at protease codon 82 in HIV-1 subtype G compared to subtype B infected individuals including a novel I82M resistance mutations. *Antivir Ther* 2005; **10**:S151.
30. Mohri H, Singh MK, Ching WT, Ho DD. Quantitation of zidovudine-resistant human immunodeficiency virus type 1 in the blood of treated and untreated patients. *Proc Natl Acad Sci U S A* 1993; **90**:25–29.
31. Albert J, Wahlberg J, Leitner T, Escanilla D, Uhlen M. Analysis of a rape case by direct sequencing of the human immunodeficiency virus type 1 pol and gag genes. *J Virol* 1994; **68**:5918–5924.
32. De Meyer S, Vangeneugden T, Lefebvre E, *et al.* Phenotypic and genotypic determination of resistance to TMC114: pooled analysis of POWER 1, 2, and 3. *Antivir Ther* 2006; **11**:S83.
33. Abecasis AB, Deforche K, Snoeck J, *et al.* Protease mutation M89I/V is linked to therapy failure in patients infected with the HIV-1 non-B subtypes C, F or G. *AIDS* 2005; **19**:1799–1806.
34. Kantor R, Katzenstein DA, Efron B, *et al.* Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: results of a global collaboration. *PLoS Med* 2005; **2**:e112.
35. Brenner B, Turner D, Oliveira M, *et al.* A V106M mutation in HIV-1 clade C viruses exposed to efavirenz confers cross-resistance to non-nucleoside reverse transcriptase inhibitors. *AIDS* 2003; **17**:F1–F5.
36. Grossman Z, Istomin V, Averbuch D, *et al.* Genetic variation at NNRTI resistance-associated positions in patients infected with HIV-1 subtype C. *AIDS* 2004; **18**:909–915.
37. Cane PA, de Ruiter A, Rice P, Wiselka M, Fox R, Pillay D. Resistance-associated mutations in the human immunodeficiency virus type 1 subtype c protease gene from treated and untreated patients in the United Kingdom. *J Clin Microbiol* 2001; **39**:2652–2654.
38. Sugiura W, Matsuda Z, Yokomaku Y, *et al.* Interference between D30N and L90M in selection and development of protease inhibitor-resistant human immunodeficiency virus type 1. *Antimicrob Agents Chemother* 2002; **46**:708–715.
39. Grossman Z, Paxinos EE, Averbuch D, *et al.* Mutation D30N is not preferentially selected by human immunodeficiency virus type 1 subtype C in the development of resistance to nelfinavir. *Antimicrob Agents Chemother* 2004; **48**:2159–2165.

40. Calazans A, Brindeiro R, Brindeiro P, *et al.* Low accumulation of L90M in protease from subtype F HIV-1 with resistance to protease inhibitors is caused by the L89M polymorphism. *J Infect Dis* 2005; **191**:1961–1970.
41. Gifford R, de Oliveira T, Rambaut A, *et al.* Assessment of automated genotyping protocols as tools for surveillance of HIV-1 genetic diversity. *AIDS* 2006; **20**:1521–1529.
42. Lindstrom A, Ohlis A, Huigen M, *et al.* HIV-1 transmission cluster with M41L 'singleton' mutation and decreased transmission of resistance in newly diagnosed Swedish homosexual men. *Antivir Ther* 2006; **11**:1031–1039.
43. Carr JK, Nadai Y, Eyzaguirre L, *et al.* Outbreak of a West African recombinant of HIV-1 in Tashkent, Uzbekistan. *J Acquir Immune Defic Syndr* 2005; **39**:570–575.

Accepted for publication 13 January 2008