Published in final edited form as:

J Acquir Immune Defic Syndr. 2018 September 01; 79(1): e21–e29. doi:10.1097/QAI.0000000000001776.

Impact of Changes over Time in the Stanford University Genotypic Resistance Interpretation Algorithm

SA Hart, Ph.D.,

Frontier Science Foundation, Amherst, NY, USA

S Vardhanabhuti, Ph.D.,

Harvard T.H. Chan School of Public Health, Boston, MA, USA

SA Strobino, B.S., and

Frontier Science Foundation, Amherst, NY, USA

LJ Harrison, M.Sc.

Harvard T.H. Chan School of Public Health, Boston, MA, USA,

Abstract

Introduction—The Stanford HIV-1 genotypic resistance interpretation algorithm has changed substantially over its lifetime. In many studies the algorithm version used is not specified. It is easy to assume results across versions are comparable, but the effects of version changes on resistance calls are unknown. We evaluate these effects for 20 antiretroviral drugs.

Methods—We calculated resistance interpretations for the same 5,993 HIV-1 sequences, from participants in AIDS Clinical Trials Group studies, under 14 versions of the Stanford algorithm from 2002 to 2017. Trends over time were assessed using repeated measures logistic regression. Changes in rule structure and scoring were examined.

Results—For most drugs, the proportion of high-level resistance calls on the same sequences was greater using more recent algorithm versions: 16/20 drugs showed significant upward trends. Some drugs, especially tenofovir, had a substantial increase. Only darunavir had a decrease. Algorithm changes impacted calls for subtype C more than B. For intermediate and high-level resistance combined, effects were weaker and more varied. Over time, rules in the Stanford algorithm have become more complex and contain more subrules. The types of rule changes responsible for trends varied widely by drug.

Discussion—Reporting the Stanford algorithm version used for resistance analysis is strongly recommended. Caution should be used when comparing results between studies unless the same version of the algorithm was used. Comparisons using different Stanford versions may be valid for drugs with few changes over time but for most comparisons version matters and for some drugs the impact is large.

Contact information for corresponding author hart@fstrf.org, Phone, 716.834.0900, Fax 716.834.8432.

CONFLICTS OF INTEREST

None for any of the authors

Keywords

HIV-1; Drug resistance; Stanford algorithm

Introduction

The HIV-1 genotypic resistance interpretation algorithm at Stanford University¹ has been widely used for assessing antiretroviral (ARV) drug resistance from HIV-1 sequences or mutation lists. In addition to its high scientific quality, it has practical advantages: it is free, public, transparent, and can be used interactively or automated through a web service.

The algorithm has evolved over time. Over 60 versions have been released since its inception in 2002. Each update changes the rules relating mutational patterns to drug resistance, based on new research findings. When analyzing data on the Stanford website, the version current at the time of analysis is used. Consequently, published drug resistance frequencies stem from varying versions.

Differences among algorithms have been studied mostly by comparing results from different algorithms (Rega,² ANRS,³ Stanford, IAS-USA⁴). Some reports (Snoek et al. 2006⁵, Liu et al. 2008⁶) conclude that the results are similar enough that one can confidently use any of the algorithms. Other reports (Vergne et al. 2006⁷, Assoumou et al. 2008⁸) point to frequent discordance.

Fewer studies have focused on version changes *within* the same algorithm. Green et al.⁹ compared results on the same sequences using various versions of the IAS-USA mutation lists from 2000 to 2006, finding substantial differences in results but no trend over time. Wagner et al.¹⁰ compared Stanford, ANRS, and Rega in versions from 2004 and 2013, focusing however on differences among algorithms, and comparing versions only for all three algorithms combined. In that pooled comparison, the authors found an increase from 55% showing resistance to at least one drug when using the 2004 versions to 62% using the 2013 versions, for the same sequences.

Here we report results from a cross-protocol analysis of AIDS Clinical Trials Group (ACTG) sequencing data, designed to evaluate whether different Stanford algorithm versions produce systematically different results, and if so, for which drugs, to what degree, and with what trends over time.

Methods

The Stanford algorithm evaluates resistance for any drug based on mutations in the relevant gene, using rules stated in an XML file. The rule evaluating resistance for a drug contains subrules, each of which assigns points if a predicate condition is met. The condition can be simple (true if a particular mutation is present) or complex (e.g., true if 3 specified mutations are simultaneously present). Points are added up across all subrules for a drug, yielding a raw score. This score is mapped, using an equivalence table embedded in the XML file, to one of five Stanford resistance levels: susceptible (<10 points), potential low-level resistance

(10–14), low-level resistance (15–29), intermediate resistance (30–59), or high-level resistance (60+). The names of the levels, but not their meaning or use within the algorithm, have changed slightly over time.

The open-source resistance calculation technology used in this paper was developed by Frontier Science Foundation (FSF) and has been used as a resistance calculation engine by the Stanford website.¹¹ It allows calculation of resistance calls for any sequence under any version of any algorithm, provided that the rules are stated in an XML file following a formal but English-like grammar. The technology compiles the algorithm and immediately executes it for input sequences.

Using this technology, we analyzed 5,993 HIV-1 sequences, all based on plasma specimens, under 14 versions of the Stanford algorithm. We used the version released closest to the midpoint of each year from 2002 to 2017, except for 2004, for which no algorithm XML file was available, and 2015, when no version was released. The specific versions are shown in the top row of Table 1.

We analyzed one HIV-1 *pol* sequence per participant in all 50 ACTG studies (19 recruiting ARV-naïve participants, 26 experienced, and 5 both) that had conducted sequencing by July 2015. Too few sequences for useful analysis included integrase; this region was therefore not assessed. Sequences failing quality control tests standardly used on ACTG sequences, or where data required for sampling was missing, were excluded. For participants sequenced more than once, preference was given to sequences from specimens collected when a participant was drug-naïve.

The ARV drugs present in each algorithm version varied. The analysis included all protease inhibitor (PI), nucleoside reverse transcriptase inhibitor (NRTI), and non-nucleoside reverse transcriptase inhibitor (NNRTI) drugs found in any version, except for amprenavir, ritonavir, and zalcitabine, which were present in only a few early versions.

Each sequence was subtyped by its genetic distance, calculated using the PHYLIP dnadist program,¹² from ~150 reference sequences from Los Alamos National Laboratory, representing pure subtypes and circulating recombinants.¹³

Two binary outcome variables were examined for each sequence under each algorithm version: high-level resistance vs. all other levels; and high-level *or* intermediate resistance vs. susceptible, potential low-level, or low-level resistance. Repeated measures logistic regression via the generalized estimating equations method was used to assess trends of resistance levels over algorithm years (by assessing algorithm year as covariate) and also whether these trends varied by HIV-1 subtype. The trends that reached statistical significance were considered “weak” when the odds ratio (OR) was between 0.97 and 1.03. For a sensitivity analysis, we created an alternative sample choosing the latest sequence for participants with more than one.

We undertook two explorations of the mechanisms behind changes in algorithm results. First, we examined the “density” of the rules in each version, for each of the 12 drugs found

in all ¹⁴ versions, measuring the mean number of simple and complex subrules, codon positions taken into account, and mutations scored by the rules.

Second, we did case studies of changes between two years for 4 drugs. We studied tenofovir (TDF) and stavudine (d4T), which had the strongest increase among reverse transcriptase inhibitors, from 2002 to their peak resistance calls in 2014; darunavir (DRV) from 2006 (when it entered the Stanford algorithm) to 2017 as the one drug with a clear downward trend; and rilpivirine (RPV) from 2014 to 2016 as an exceptionally large short-term increase.

For each subrule that yielded points for a particular drug on the sequence under analysis, we determined the predicate condition, the points assigned because it was met, and the mutation(s) that activated it. This allowed us to identify the rule changes principally responsible for the aggregate changes between the two years. This data, because of the logic of Stanford algorithms, was at the level of the subrule, not the mutation. Some mutations appear in more than one subrule; subrules can score more than one amino acid mutation and/or mutations in more than one codon. Therefore, the attribution of points to mutations was approximate. In this analysis, we treated thymidine analogue mutations (TAMs) as a single unit.

Results

Of the 5,993 HIV-1 sequences sampled, 2,641 (44%) were ARV-naïve, 3,135 (52%) experienced and 217 (4%) unknown. Overall, 2,660 (44%) of the sequences were subtype B virus and 1,690 (28%) subtype C, with the rest other subtypes or unclassifiable. The specimens dated from 1992 to 2015.

High-level Resistance Changes

Table 1 displays changes over time in calls of high-level resistance. The patterns varied by drug, but the most common pattern was an increase in the proportion of sequences with calls of high-level resistance. There were statistically significant upward trends for 16 of 20 drugs, 8 of them strong; a downward trend for DRV; and no trend for the other three. Figure 1 graphs these results.

All PIs except DRV had a statistically significant upward trend. At the top end, high-level resistance calls for lopinavir (LPV) were about 2.7 times as likely under the 2017 as under the 2002 version. Fosamprenavir and atazanavir resistance calls increased about 50%.

Four of the seven NRTIs had a statistically significant upward trend. TDF had the strongest, with the proportion receiving high-resistance calls 14-fold higher under the 2017 version than the 2002 version. Didanosine and d4T showed increases of 138% and 71%, respectively, while zidovudine had a significant but weak trend.

All the NNRTIs had significant upward trends. The only strong trends were for RPV (171% increase) and etravirine (ETR; 68% increase).

The tests for interaction by HIV-1 subtype are shown in the right-hand panel of Table 1. Nine of the 20 drugs had statistically significant interactions, 5 of them strong. All of the

significant ORs were >1 . Except for DRV, this means that the upward trend was stronger for subtype C than for B. For DRV, the trend was downward for subtype B; for C, essentially flat.

The results from the alternative sample were very similar to those from the main one. The ORs measuring trends were statistically indistinguishable for all 20 drugs.

In addition to the trends, the results show year-to-year variations, sometimes in the opposite direction of the overall trend for the drug.

Intermediate or High-level Resistance Changes

Table 2 displays the changes in calls of intermediate and high-level resistance combined. The trends for this outcome were weaker and more varied than for calls of high-level resistance alone. Twelve drugs showed an upward trend, two no significant trend, and the other 6 a downward trend. Even the significant trends were modest; only one downward and one upward trend were strong, and the largest proportionate increase from 2002 to 2017 was 40%, versus 14-fold for high-level resistance.

Five PIs showed weak upward trends, one no change, and two downward trends. The most striking trend was for DRV, with only 5% called intermediate or high resistance in 2017, vs. 13% in 2006, when the drug first appeared in the algorithm.

Three NRTIs showed statistically significant but tiny ($OR \sim 1.00$) downward trends, one no trend, and three weak upward trends; the largest were for d4T, from 18% to 25%, and TDF, from 12% to 16%.

Of the NNRTIs, only RPV had a strong trend; ETR showed a weak downward trend and the others upward trends.

Subtype interactions followed no clear pattern: only 7 drugs showed a statistically significant interaction, and the effects varied in direction, with 2 drugs having $ORs < 1$. The interaction effect was strong only for TDF and (in the opposite direction) DRV.

The results from the alternative sample were statistically indistinguishable from those for the main one.

Rule Density

The bottom rows in Table 2 show the means, for the 12 drugs in all 14 algorithm versions, on the density measures. Figure 2 graphs the data. The mean number of distinct codon positions scored per drug did not increase, and the number of distinct scored mutations, while changing considerably from version to version, did not show a statistically significant trend (Spearman correlation 0.23, $p = .43$). What did increase was the mean number of subrules (Spearman correlation 0.88, $p < .0001$), from 19.4 in 2002 to 28.5 in 2017. Algorithm versions from 2012 on, in particular, had many more subrules than previous versions. However, this increase was entirely due to an increase in complex subrules (e.g. “67EGN AND 70R AND 219ENQR”). These new multi-codon rules, added to previously existing rules that were usually not removed or given reduced points, were part of how

scores, and therefore resistance calls, increased over time. In short, the algorithm evolved from rules solely focusing on point mutations toward concern with interactions among mutations in different codons.

Case Studies of Algorithm Change

Table 3 shows results concerning 4 drugs. The analysis was performed on the “trend cases”: sequences that had a call of high-level resistance in the year with more such calls (the later year, except for DRV) and a lower level in the other year.

The upper part of the table describes overall patterns, which vary greatly by drug. For TDF and d4T, new subrules were the most important factor driving increased resistance calls. Changes in rules (more points and/or more scored amino acid mutations) for the same codon(s) were important for d4T and made almost all the difference for RPV. Dropping older rules was the key factor in the decrease of DRV resistance calls but made no difference for other drugs.

The mean points awarded by each subrule, across sequences where that subrule was triggered, went up for all four drugs, suggesting a general trend toward assigning higher point values. Complex subrules played a critical role for TDF and were important for d4T but not for RPV or DRV.

Lastly, the changes in d4T, and to some extent TDF, were driven by major scoring changes, where a new or revised subrule yielded a score increase of 20 or more points for each affected sequence. For DRV, at the other extreme, the decrease was driven almost entirely by changes affecting the score by less than 10 points, such as dropping old rules that had yielded a few points each. For RPV and TDF mid-range scoring changes, yielding 10–19 point increases, had the most effect.

The lower part of the table shows the particular mutations that had the most important effect on the trend for each drug. The focus here is on change; for example, RT69 insertions yield high point values for NRTIs but are not shown because their scoring did not change much. The mutations vary in their prevalence (the first numeric column) and the magnitude of their effects (the next two columns); the final column shows how many individual sequences were pushed across the boundary between high-level and lower levels of resistance by the mutation.

For TDF, the high-impact changes concerned TAMs and K65R. For d4T, rescoring of TAMs was the major factor driving the trend. For RPV, the change in Y181C scoring, while only adding 15 points, affected two-thirds of the trend cases and by itself was enough to push one-third of them from lower levels to high resistance. M230L also played a role for RPV. In the downward trend for DRV, L90M, while only changing from 8 points to not being scored, was found in almost all trend sequences, and by itself was enough to move the resistance call down for 71% of them.

DISCUSSION

The Stanford algorithm has changed over time. For most drugs, there was a significant upward trend in resistance calls from older to newer versions, with TDF manifesting the strongest trend. This pattern was more pronounced for calls of high-level resistance than for calls of intermediate or high-level resistance combined; thus, much of the impact of algorithm changes was in moving sequences from intermediate to high-level resistance, rather than from lower levels to intermediate resistance. One possible explanation is the increasing use of multi-codon subrules. These are triggered only when there are resistance mutations at more than one codon, and therefore more often by sequences showing significant resistance than by lower-scoring ones.

For some drugs, especially for high-level resistance calls, algorithm changes had a more pronounced effect on C than B sequences. One possible explanation is that each algorithm version is based on the currently existing body of resistance research results, and results for subtype C were sparser in 2002 than in 2017.

Why do we see the general tendency for resistance calls to rise as the Stanford algorithm gets revised?

The algorithm is based on systematic analysis of the results of resistance research, and the body of results grows continuously. New resistance patterns are discovered and get added to the algorithm. In addition, publication bias (see for example Easterbrook et al. 1991¹⁴) may conceivably reduce the chances of our learning about non-replication of previously reported patterns. Each time a new association is published, the algorithm has a chance of acquiring a new subrule, or more points for existing subrules, for the drug in question, and less often loses an existing subrule. Thus over time the number of points the same sequence accrues for a drug is more likely to increase than decrease. Evidence consonant with this hypothesis includes the trend toward more subrules per drug, and (in the case studies) the increase in the mean number of points conferred by active subrules.

A tendency toward point escalation would not necessarily have led to more resistance calls. After calculating raw points, Stanford algorithms convert these to resistance levels using an equivalence table that could be altered when creating a new version. One could, for example, set the threshold for high-level resistance at 75 points instead of 60. However, the equivalence table has not been changed during the life of the Stanford algorithm, and therefore the increase over time in points, for most drugs, has led to increases in resistance calls. The patients have obviously not developed new resistance because of changes in the algorithm, but one might interpret the upward trend in resistance calls as saying that their HIV drug resistance is worse than we previously realized.

Limitations

ACTG protocol participants do not represent the overall population of HIV-1 infected individuals. While we did an internal sensitivity analysis based on an alternative sample, generalizability to other populations is not guaranteed. Furthermore, the majority of the sequences from participants who were ARV-naïve at the time of blood draw were from

studies that had entry requirements excluding participants with certain pre-existing resistance or mutation patterns. Another limitation is that subtypes other than B and C were poorly represented, making it impossible to do a broader analysis of how changes in the Stanford algorithm may differentially affect subtypes. Finally, integrase inhibitors, which entered the Stanford algorithm in 2009, could not be studied because the ACTG had not yet performed large numbers of integrase sequencing when the data for the analysis was finalized. With regard to the case studies of rule changes, we do not know how the findings would look for other drugs.

The analysis in this paper is addressed primarily to research and epidemiological use of resistance results, and especially to trend and comparative analysis, rather than to clinical practice. In a clinical context, version comparisons are normally unimportant, and it may make sense to use the most recent version of whatever algorithm one chooses, since that version will reflect the most complete set of research findings.

Conclusions

Comparisons of drug resistance levels across populations or over time can be confounded with differences in algorithm versions. This issue arises directly when performing meta-analyses of drug resistance results, as in Gupta et al.,¹⁵ or proposing use of the Stanford algorithm in multiple future studies, as in part of the research suggested by Jordan et al.¹⁶

The broader problem is informal comparisons. For example, one study compared Brazilian findings on resistance prevalence to “what has been described in Europe and the United States,”¹⁷ and another compared its results to those from nine different studies from various parts of the world.¹⁸ The problem is that the authors making the comparisons either did not know or did not tell readers which specific algorithm and version was used in those other studies, and without that knowledge one cannot say whether the comparisons are informative.

In addition to trends, the year-to-year variations in results, which are probably driven partly by the particular moment when Stanford updated the algorithm for a group of drugs, incorporating a bolus of new findings, are sometimes large enough to potentially affect longitudinal or intergroup comparisons.

We recommend that all published or publicly presented resistance results generated using the Stanford algorithm state the exact version used. If the analysis was completed online and the version is not known, the date on which the analysis was completed and the specific URL used (since the Stanford website has several different systems) should be stated. Similar principles should be followed when using other algorithms.

Also, caution should be used when performing meta-analyses or comparing published results from different populations or time periods. If the version used for either dataset is not known, comparisons cannot be relied upon unless the numbers for the drug in Tables 1 and 2 are essentially constant across versions. If the versions are known but differ, comparisons are reasonable if the numbers are very close for the years of the two versions.

When comparing results from different studies or populations, the preferred approach is to obtain the raw sequencing data and reanalyze it so that all results stem from the same algorithm version. One option is to run all the sequences at once on the Stanford website. Another, if the raw data is available for only one of the studies or a specific version is required, is to use the technology on which this paper is based.¹¹

We found substantial changes in the proportion of resistance calls, for the exact same sequences, for most of the ARVs drugs studied. The dominant tendency was for resistance calls to increase over time. This applied to 16/20 of the drugs for high-level resistance, and 12/20 for high-level or intermediate resistance combined. However, we observed great variation in how much a given drug was affected by version changes, ranging from a 14-fold increase in calls of high-level resistance for TDF to a decrease for DRV. In addition to secular trends, there are year-to-year differences for most drugs. In conclusion, algorithm version matters, significantly affecting the results scientists report, and close attention to its impact is essential for accumulating and comparing HIV-1 drug resistance results across different populations and over time.

Acknowledgements

In addition to the funding and support from the NIH and ACTG, we wish to thank colleagues who contributed to the work, including Laura Hovind and Adina Stoica of FSF, who were instrumental in the development and testing of the software used in the project; Konstantia Angelidou and Ronald Bosch of the Harvard T. H. Chan School of Public Health for input on this research; and Robert Shafer, who has shared all the versions of the Stanford algorithm with us and been a continuous collaborator for over a decade. Finally, we thank the ACTG protocols teams that originally collected the sequence data, the sites and laboratories involved in the sequencing work, and the participants in the studies.

Funding

This research was supported, as part of the AIDS Clinical Trials Group, by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award numbers UM1 AI068634, UM1 AI068636 and UM1 AI106701. The content is solely the responsibility of the authors and does not necessarily represent the views of the National Institutes of Health.

References

1. The Stanford web site. Available at: <http://hivdb.stanford.edu>. Accessed June 30, 2018.
2. Rega algorithms. Available at: https://rega.kuleuven.be/cev/avd/files/software/rega_algorithm. Accessed Jun 30, 2018.
3. The latest version of ANRS algorithm can be found at <http://www.hivfrenchresistance.org/2016/Algo2016-hiv1.pdf>
4. The most recent IAS-USA mutation lists and commentary can be found at <https://www.iasusa.org/sites/default/files/uploads/2017hiv-muta-article.pdf> or in print: Annemarie M Wensing A, Calvez V et al. 2017 Update of the Drug Resistance Mutations in HIV-1. Topics in Antiviral Med: 2017:24(4)
5. Snoeck J, Kantor R, Shafer R. Discordances between interpretation algorithms for genotypic resistance to protease and reverse transcriptase inhibitors of human immunodeficiency virus are subtype dependent. Antimicrob Agents Chemother. 2006;50:694–701. [PubMed: 16436728]
6. Liu L, May S, Richman D, et al. Comparison of algorithms that interpret genotypic HIV-1 drug resistance to determine the prevalence of transmitted drug resistance. AIDS. 2008;22(7):835–839. [PubMed: 18427201]
7. Vergne L, Snoeck J, Aghokeng A, et al. Genotypic drug resistance interpretation algorithms display high levels of discordance when applied to non-B strains from HIV-1 naive and treated patients. FEMS Immuno Med Micro. 2006;46:53–62.

8. Assoumou L, Brun-Vézinet F, Cozzi-Lepri A, et al. Initiatives for Developing and Comparing Genotype Interpretation Systems: External Validation of Existing Systems for Didanosine against Virological Response. *J Infect Dis.* 2008;198:470–80. [PubMed: 18598191]
9. Green H, Tilston P, Fearnhill E, et al. The impact of different definitions on the estimated rate of transmitted HIV drug resistance in the United Kingdom. *J Acquir Immune Defic Syndr.* 2008; 49(2):196–204. [PubMed: 18769348]
10. Wagner S, Kurz M, Klimkait, et al. Algorithm evolution for drug resistance prediction: comparison of systems for HIV-1 genotyping. *Antiviral Therapy.* 2015;20:661–665. [PubMed: 25710167]
11. The software developed by FSF for compiling and executing any XML-defined algorithm is now open-source and can be accessed at https://github.com/FrontierScience/asi_interpreter. Guidance for using this software can be obtained by e-mailing seqdev@fstfrf.org.
12. From University of Washington, available at <http://evolution.genetics.washington.edu/phylip/>
13. Los Alamos National Laboratory HIV-1 reference sequences for subtyping can be downloaded from <https://www.hiv.lanl.gov/content/sequence/NEWALIGN/align.html>.
14. Easterbrook PJ, Gopalan R, Berlin JA, et al. Publication bias in clinical research. *Lancet.* 1991; 337(8746):867–872. [PubMed: 1672966]
15. Gupta R, Jordan M, Sultan B, et al. Global trends in antiretroviral resistance in treatment-naïve individuals with HIV after rollout of antiretroviral therapy treatment in resource-limited settings: a global collaborative study and meta-regression analysis. *Lancet.* 2012;380(9849):1250–158. [PubMed: 22828485]
16. Jordan M, et al. World Health Organization surveys to monitor HIV drug resistance prevention and associated factors in sentinel antiretroviral treatment sites. *Antiviral Therapy.* 2008;13 Supp 2:15–23.
17. Barreto C, Nishyia A, Araújo L, et al. Trends in Antiretroviral Drug Resistance and Clade Distributions Among HIV-1--Infected Blood Donors in Sao Paulo, Brazil. *J Acquir Immune Defic Syndr.* 2006;41(3):338–341. [PubMed: 16540943]
18. Masquelier B, Bhaskaran K, Pillay D, et al. Prevalence of Transmitted HIV-1 Drug Resistance and the Role of Resistance Algorithms: Data From Seroconverters in the CASCADE Collaboration From 1987 to 2003. *J Acquir Immune Defic Syndr.* 15 12 2005;40(5):505–511. [PubMed: 16284524]

Figure 1.

Calls of high-level resistance, by algorithm year, for darunavir and drugs with odds ratio > 1.03 or < 0.97

Figure 2.
Rule density measures

Table 1.

Algorithm version and calls of high-level resistance by algorithm year and drug

Year	Algorithm version & percent called as high-level resistance for each algorithm year ^b														Trend ^a			Subtype interaction				
	Version	2002	2003	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2016	2017	Direction and fold change	Odds ratio and confidence interval	p-value	Odds ratio	p-value		
		2002.03	2003.08	4.1.3	4.2.1	4.2.9	4.3.6	6.0.2	6.0.9	6.0.12	6.2	6.3	7.0.1	8.1.1	8.4							
Drug																						
PIs	LPV	4.9%	5.1%	6.7%	7.5%	6.0%	6.0%	7.9%	7.9%	7.9%	10.1%	10.1%	12.9%	13.1%	13.1%	up	2.7	1.07 (1.06–1.08)	<.001	1.02	0.187	
	FPV				9.4%	8.5%	8.5%	10.2%	10.2%	10.2%	13.4%	13.4%	13.6%	14.0%	14.0%	up	1.5	1.04 (1.03–1.04)	<.001	1.03	0.015	
	ATV		9.6%	10.4%	10.2%	9.4%	9.6%	11.0%	11.0%	11.0%	13.2%	13.2%	14.2%	14.2%	14.4%	up	1.5	1.03 (1.03–1.04)	<.001	1.04	<.001	
	SQV	10.2%	10.2%	9.9%	10.0%	9.6%	9.6%	9.8%	9.8%	9.8%	11.7%	11.7%	12.0%	13.7%	13.7%	weak up	1.3	1.03 (1.02–1.03)	<.001	1.07	<.001	
	IDV	11.7%	11.6%	11.4%	11.7%	11.1%	11.1%	12.1%	12.2%	12.1%	12.6%	12.6%	14.3%	14.3%	14.3%	weak up	1.2	1.01 (1.01–1.02)	<.001	1.01	0.214	
	NFV	15.2%	15.1%	15.1%	15.2%	14.6%	14.9%	16.0%	16.0%	16.0%	16.3%	16.3%	16.7%	16.9%	16.9%	weak up	1.1	1.01 (1.01–1.01)	<.001	1.02	<.001	
	TPV				5.6%	3.4%	3.1%	2.5%	1.3%	1.3%	2.5%	2.5%	5.5%	5.8%	5.8%	weak up	1.0	1.02 (1.01–1.03)	<.001	1.03	0.197	
	DRV				3.0%	1.1%	0.9%	1.2%	0.5%	0.5%	1.6%	1.6%	1.8%	2.2%	2.2%	weak down	0.7	0.98 (0.97–1.00)	0.024	1.10	0.040	
	NRTIs	TDF	0.6%	0.5%	0.9%	0.8%	0.9%	1.0%	0.8%	1.8%	1.8%	6.3%	10.9%	11.3%	8.4%	8.4%	up	13.6	1.13 (1.12–1.14)	<.001	0.99	0.579
		d4T	7.4%	9.2%	10.1%	11.9%	11.9%	12.0%	12.0%	12.0%	12.0%	17.8%	18.6%	18.6%	17.6%	17.6%	up	2.4	1.05 (1.05–1.06)	<.001	1.01	0.396
ddI		10.6%	10.7%	11.8%	10.5%	10.6%	11.1%	11.0%	12.9%	12.9%	15.3%	17.7%	18.3%	18.2%	18.2%	up	1.7	1.04 (1.04–1.05)	<.001	1.01	0.173	
ZDV		11.4%	12.6%	13.3%	13.2%	13.1%	13.1%	13.0%	13.0%	13.0%	17.8%	18.8%	18.8%	16.5%	16.5%	weak up	1.4	1.02 (1.01–1.02)	<.001	0.99	0.083	
ABC		14.8%	14.8%	10.9%	10.5%	10.6%	10.8%	10.8%	12.7%	12.7%	14.4%	16.3%	17.5%	16.0%	18.5%	none	1.2	1.00 (1.00–1.01)	0.193	1.02	0.003	
NNRTIs	3TC	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	none	1.0	1.00 (1.00–1.00)	0.165	1.00	0.318	
	FTC			27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	27.1%	none	1.0	1.00 (1.00–1.00)	0.317	1.00	0.318	
	RPV										6.3%	6.9%	9.4%	13.8%	13.8%	up	2.2	1.20 (1.18–1.22)	<.001	1.06	0.005	
	ETR						2.6%	2.9%	3.8%	3.8%	3.5%	3.0%	3.8%	4.5%	4.5%	up	1.7	1.06 (1.05–1.08)	<.001	1.00	0.856	
	DLV	25.3%	26.4%	26.8%	26.9%	26.9%	27.0%	27.2%	27.2%	27.2%	31.7%	31.7%	31.8%	31.9%	31.9%	weak up	1.1	1.01 (1.01–1.01)	<.001	1.01	<.001	
EFVs	NVP	30.0%	31.5%	31.7%	31.7%	31.7%	31.7%	31.7%	31.7%	31.7%	31.7%	31.7%	31.8%	31.9%	31.9%	weak up	1.1	1.01 (1.00–1.01)	<.001	1.00	0.266	
	EFV	25.6%	25.8%	25.1%	25.1%	25.1%	25.4%	25.9%	25.9%	25.9%	25.6%	26.5%	27.0%	27.2%	27.2%	weak up	1.1	1.00 (1.00–1.01)	<.001	1.01	0.009	

Drugs: LPV=lopinavir; FPV=fosamprenavir; ATV=atazanavir; SQV=saquinavir; IDV=indinavir; NFV=nelfinavir; TPV=tipranavir; DRV=darunavir; TDF=tenofovir; ddI=didanosine; ZDV=zidovudine; ABC=abacavir; 3TC=lamivudine; FTC=emtricitabine; RPV=rilpivirine; ETR=etravirine; DLV=delavirdine; NVP=nevirapine; EFV=efavirenz

Notes: General: numbers above are rounded; statements in text are based on actual values.

^aUp/down mark statistically significant trends; “weak” means 1.03 > OR > 0.97.

^bBlank cells mean the drug was not included in that algorithm version.

c, Odds ratio (OR) parameters measure the change per year.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Drugs: LPV=lopinavir; FPV=fosamprenavir; ATV=atazanavir; SQV=saquinavir; IDV=indinavir; NFV=nelfinavir; TPV=tipranavir; DRV=darunavir; TDF=tenofovir; ddI=didanosine; ZDV=zidovudine; ABC=abacavir; 3TC=lamivudine; FTC=entriictabine; RPV=rilpivirine; ETR=etravirine; DLV=delavirdine; NVP=nevirapine; EFV=efavirenz

Notes: *General:* numbers above are rounded; statements in text are based on actual values.

^aUp/down mark statistically significant trends; “weak” means $1.03 > OR > 0.97$.

^bBlank cells mean the drug was not included in that algorithm version.

^cOdds ratio (OR) parameters measure the change per year.

Table 3.

Analysis of the impact of rule changes, for four drugs

Overall Results for Each Drug								
Drug	TDF		d4T		RPV		DRV	
Algorithm version years	Earlier	Later	Earlier	Later	Earlier	Later	Earlier	Later
	2002	2014	2002	2014	2014	2016	2006	2017
N cases called high by this version, not by other	1	592 ^a	0	617 ^a	0	241 ^a	63 ^a	19
Mean point score	36.9	76.7	42.9	91.1	46.6	64.1	65.1	34.6
Mean change in points among trend cases ^a	39.8		48.3		17.5		−30.5	
Mean change, and % of total change, due to:								
New subrules ^b	33.6 (84%)		24.3 (50%)		0.6 (1%)		4.1 (13%)	
Changed subrules ^b	8.2 (21%)		23.4 (48%)		16.9 (97%)		4.3 (14%)	
Old subrules ^b	−2.0 (−5%)		0.6 (3%)		−0.0 (−0%)		38.8 (−127%)	
Number of subrules that yielded points ^d	5.3	8.4	4.8	6.8	2.2	2.3	11.6	3.8
Mean points per subrule yielding points	7.0	9.1	8.9	13.4	21.2	27.9	5.6	9.1
Mean point change, and % of total change, due to complex subrules ^c	32.1 (81%)		20.1 (42%)		0.6 (3%)		4.0 (13%)	
Percent of change for drug from scoring shifts of:								
30+ points	20%		34%		0%		0%	
20–29 points	1%		47%		0%		0%	
10–19 points	75%		20%		85%		1%	
1–9 points	18%		10%		15%		130%	
<0 points (change in counter-trend direction)	−13%		−10%		−0%		−31%	
High-Impact Scoring Changes ^e								
Drug	Muta-tion	N cases	Maximum point gain	Mean point gain from mutation where present	Percent of total gain for drug due to this mutation ^c		N cases and percent of trend cases where enough to change call ^f	
TDF	TAMs ^g	465	60	29.7	59%		409 (69%)	
TDF	K65R	154	70	43.7	29%		154 (26%)	
d4T	TAMs ^g	592	86	42.3	84%		542 (88%)	
d4T	K65R	40	55	52.0	7%		35 (6%)	
RPV	Y181C	161	15	15.0	61%		87 (34%)	
RPV	M230L	26	45	16.1	9%		26 (11%)	

Overall Results for Each Drug						
Drug	TDF		d4T		RPV	DRV
DRV	L90M	58	-8	-8.0	24%	45 (71%)

Notes: Drugs are tenofovir (TDF), stavudine (d4T), rilpivirine (RPV), and darunavir (DRV).

^a“Trend cases”; means are on these cases unless otherwise specified.

^b“New” = for codons (or combinations of codons) that were not scored in the earlier version but were in the later version. “Changed” = where the same codon(s) was scored in both versions, but changed, for example with new non-wild-type amino acids scored, or point values altered. “Old” = existed in the earlier version, while the later version included no subrule for the codon(s) in question.

^cDuring calculations, for each sequence the predicate condition for each subrule is evaluated as true or false, and if true the number of points specified in the subrule are added to the total for the drug/sequence in question.

^dThe percentages of total change due to various factors add up to > 100%, since many cases are affected by more than one.

^eThe data available was on the scores resulting from subrules, not mutations. Since a given mutation can be referenced in more than one subrule, and in some subrules confers points only in combination with another mutation or in the absence of another mutation, the attribution of points to mutations is approximate.

^fThat is, for sequences not called as high resistance in the earlier version to reach 60 points and therefore be called high, or for DRV, to move the sequence down from a high call.

^gThe presence of any TAM (41L, 67N, 70R, 210W, 215YF, or 219QE) is reported on this line; thus, this shows the combined effect of 6 related mutations, not (as for other rows) a single mutation.