HIV-GRADE: A Publicly Available, Rules-Based Drug Resistance Interpretation Algorithm Integrating Bioinformatic Knowledge

Martin Obermeier^a Alejandro Pironti^l Thomas Berg^a Patrick Braun^c
Martin Däumer^d Josef Eberle^e Robert Ehret^c Rolf Kaiser^g Niels Kleinkauf^b
Klaus Korn^h Claudia Kücherer^b Harm Müllerⁱ Christian Noah^j
Martin Stürmer^k Alexander Thielen^{d,l} Eva Wolf^f Hauke Walter^h

^aMedizinisches Labor Dr. Berg and ^bRobert Koch-Institute, Berlin, ^cPZB Aachen, Aachen, ^dInstitut für Immunologie und Genetik, Kaiserslautern, ^eMax von Pettenkofer Institute and ^fMUC Research, Munich, ^gInstitute of Virology, University of Cologne, Cologne, ^hInstitute of Virology, University of Erlangen-Nürnberg, National Reference Center for Retroviruses, Erlangen, ⁱLabor Fenner and ^jLabor Lademannbogen, Hamburg, ^kJohann Wolfgang Goethe-University Hospital, Institute for Medical Virology, Frankfurt, and ^lMax Planck Institute for Informatics, Saarbrücken, Germany

Key Words

HIV-GRADE · Genotypic drug resistance · Clinical validation

Abstract

Background: Genotypic drug resistance testing provides essential information for guiding treatment in HIV-infected patients. It may either be used for identifying patients with transmitted drug resistance or to clarify reasons for treatment failure and to check for remaining treatment options. While different approaches for the interpretation of HIV sequence information are already available, no other available rules-based systems specifically have looked into the effects of combinations of drugs. HIV-GRADE (Genotypischer Resistenz Algorithmus Deutschland) was planned as a country-wide approach to establish standardized drug resistance interpretation in Germany and also to introduce rules for estimating the influence of mutations on drug combinations. The rules for HIV-GRADE are taken from the literature, clinical follow-up data and from a bioinformatics-driven interpreta-

tion system (geno2pheno_[resistance]). HIV-GRADE presents the option of seeing the rules and results of other drug resistance algorithms for a given sequence simultaneously.

Methods: The HIV-GRADE rules-based interpretation system was developed by the members of the HIV-GRADE registered society. For continuous updates, this expert committee meets twice a year to analyze data from various sources. Besides data from clinical studies and the centers involved, published correlations for mutations with drug resistance and genotype-phenotype correlation data information from the bioinformatic models of geno2pheno are used to generate the rules for the HIV-GRADE interpretation system. A freely available online tool was developed on the basis of the Stanford HIVdb rules interpretation tool using the algorithm specification interface. Clinical validation of the interpretation system was performed on the data of treatment episodes consisting of sequence information, antiretroviral treatment and viral load, before and 3 months after treatment change. Data were analyzed using multiple linear regression. **Results:** As the developed online tool allows easy

comparison of different drug resistance interpretation systems, coefficients of determination (R^2) were compared for the freely available rules-based systems. HIV-GRADE ($R^2 = 0.40$), Stanford HIVdb ($R^2 = 0.40$), REGA algorithm ($R^2 = 0.36$) and ANRS ($R^2 = 0.35$) had a very similar performance using this multiple linear regression model. **Conclusion:** The performance of HIV-GRADE is comparable to alternative rules-based interpretation systems. While there is still room for improvement, HIV-GRADE has been made publicly available to allow access to our approach regarding the interpretation of resistance against single drugs and drug combinations.

Copyright © 2012 S. Karger AG, Basel

Introduction

The treatment failure of highly active antiretroviral therapy can be caused by a variety of factors. Besides a low adherence to the combination treatment or an insufficient uptake of the drugs, drug resistance is one of the most common reasons for the insufficient inhibition of viral replication [1]. The high mutation rate of HIV leads to a fast evolution towards drug-resistant strains [2]. Some drugs show a very specific mutation pattern and assessment can be very easy. This is true if single-point mutations like the mutation K103N in the reverse transcriptase (RT) lead to resistance against the nonnucleoside RT inhibitors efavirenz and nevirapine [3]. Other drugs, by contrast, show very complex mutation patterns like most of the protease inhibitors or the nucleosidic/nucleotidic RT inhibitors (NRTI/NtRTI) [4]. Interaction between the mutations can be very complex. Some mutations can even cause a resensitizing effect on specific drugs while leading to resistance against the drugs selecting them [5, 6]. At present, 25 drugs are approved by the FDA for the treatment of HIV. Some drugs are available as fixed-dose regimens, which underlines the need for analyses of combination effects in assessing the clinical relevance of resistance-associated mutations.

Different studies have shown an improvement in treatment outcome when treatment decisions are supported by drug resistance testing [7, 8]. However, in some studies, this benefit appears to be only short-lived [9], while in others it does not exist at all [10]. Many possible reasons for these differences have been discussed, ranging from the insufficient standardization of interpretation [11] to local treatment guidelines with preferences for specific drug combinations.

HIV-GRADE (Genotypischer Resistenz Algorithmus Deutschland) has tried to overcome some of those prob-

lems by generating drug resistance interpretation rules which are adapted to national treatment guidelines and are clinically validated on treatment-response data sets from Germany.

While rules-based analyses of drug resistance mutations can still be carried out by hand, based on published tables of rules, this is very labor intensive, time consuming and prone to errors. To overcome this problem and to make access as easy as possible, it was decided by the HIV-GRADE group at an early stage to implement a publicly available internet-based tool.

Material and Methods

The generated rules are based on different sources. Besides a classic literature search for published data on drug resistance-associated mutations, the analysis of data derived from clinical studies and case reports identified in HIV-GRADE-associated centers and data from the bioinformatic models implemented in the geno2pheno_[resistance] [12] system were made available to and were discussed by all HIV-GRADE members. The final decision on how and whether to implement a mutation in a certain rule was done by vote.

Based on the publicly available Perl source code of the 2003 version of the Stanford HIVdb tool [13], we developed the HIV-GRADE internet tool [14]. A characteristic feature of this tool is its capacity to perform side-by-side comparisons of the results of various drug resistance assessment systems. While the main workflow of the program did not change, much of the Stanford HIVdb tool was reprogrammed or completely dropped, e.g. the database analysis. The implemented workflow can be split into 3 parts: (1) extraction of mutations compared to an HIV-1 subtype B consensus sequence, (2) application of the rules on this set of mutations, and (3) output of the results. When the data are already provided as a mutation list, step 1 is skipped. Since most of the sequences entered into the system are nucleic acid sequence strings, which have not been translated into protein sequence strings, special care has to be taken in the alignment and translation of sequence information. Nucleic acid sequence strings are directly aligned to an HIV-1 subtype B consensus amino acid sequence of the target region. This HIV-1 subtype B consensus sequence is the same as that used by the Stanford HIVdb tool. Alignment is performed using the Lap-Tool [15], which is also used by the Stanford HIVdb tool. This tool, however, was modified for use in the HIV-GRADE tool by adding a supplemental gap-weighting parameter to avoid breaking the translation frame: 'in-reading-frame' gaps receive a bonus weight.

One of the major advantages of this approach is the option of using the algorithm specification interface [16], which uses an XML container. Although the algorithm specification interface only has a limited number of syntactical elements, it is quite extensive and even complicated rules can be described in this format. For most of the freely available rules-based interpretation systems, the XML files are published online.

Multiple nucleic acid sequences can be analyzed by batch with the HIV-GRADE tool. At the same time, the results of HIV-GRADE can be compared to the results of the ANRS [17], REGA

Table 1. Geno2pheno_[resistance] scores for AZT

Amino acid position of the RT	Mutation	Score
69	insertion	0.77
151	M	0.47
215	Y	0.38
215	F	0.36
70	R	0.33
210	W	0.23
103	N	-0.03
65	R	-0.06
181	C	-0.11
184	V	-0.19
184	I	-0.25

This table presents an excerpt of the scores for calculating resistance against AZT. Scores are sorted by value in a descending order. Positive scores lead to an increase in estimated resistance, while negative scores lower resistance against the drug.

[18] and HIVdb [13] rule sets and the geno2pheno_[resistance] system [19]. Geno2pheno_[resistance] is a data-based system that performs quantitative predictions on viral drug resistance by applying support vector regression to a set of matched phenotype-genotype pairs.

The batch analysis feature has been used regularly for the clinical validation of the rule set. A database consisting of 334 treatment change episodes (TCE) is the basis of this validation. The TCEs are defined by a treatment change that occurred no later than 4 weeks after a genotypic drug resistance test was performed, the drug combination taken by the patient and a 12-week follow-up with a viral load measurement.

An active drug score for each drug was computed by casting the qualitative drug-specific susceptibility levels of each algorithm to equidistant numeric values (e.g. ANRS: 'susceptible' = 1, 'possible resistance' = 0.5, 'resistance' = 0). Furthermore, a multiple linear regression model was developed, based on the following assumptions. (1) A treatment that includes many active drugs leads to a higher decline in viral load than a treatment with poor activity [20]. (2) Each drug should have its own weighting factor applied to the particular activity score defined by the rules to take into account the different potency of each drug. It has been shown that this drug potency weighting increases the accuracy of the model used [21]. (3) A high viral load at the beginning of the TCE allows a higher viral load decrease than one near the detection cutoff.

$$\log \left(\frac{VL_{after_change}}{VL_{before_change}} \right) = \sum_{v=1}^{\# \text{ drugs}} a_{drug_v} \cdot \Delta ADS_{drug_v} + b \cdot \log(VL_{before_change}) + c \quad (1)$$

This multiple linear regression model was applied to the different drug resistance algorithms.

As data on specific drug combinations was insufficient in the basic validation data, additional data were collected for the analysis of quadruple NRTI abacavir (ABC), azidothymidine (AZT), lamivudine (3TC) and tenofovir (TDF) regimens or treatment

Table 2. HIV-GRADE rules for AZT efficacy of inhibition of the RT-associated amino acid mutations

	Resistant	Intermediate resistance	Limited susceptibility
AZT	151M 69 insertion 3 TAMs	2 TAMs	1 TAM

Thymidine analog mutations (TAMs): M41L, D67N, K70R, L210W, T215Y/F, K219Q/E.

For each of 74V, 181C or 184V/I, subtract 1 from 67N, 70R or 219Q/E in the absence of 41L+210W+215Y/F. For 2 out of 74V, 181C or 184V/I, subtract 1 TAM in the presence of at least 41L+210W+215Y/F. For 3 out of 74V, 181C or 184V/I, subtract 2 TAMs in the presence of at least 41L+210W+215Y/F. For 65R subtract 2 TAMs. Maximum enhancement is one category, except for 65R. 'Hypersusceptible' if 65R without any TAMs, 151M and 69 insertion. 'Hypersusceptible' if 2 out of 184V, 74V and 181C without any TAMs, 151M and 69 insertion.

with darunavir (DRV)-containing drug combinations. While the linear regression method for clinical validation was applied to the DRV data set, simple calculation of success rates sufficed for the analysis of the quadruple NRTI regimens (since treatment is uniform). In this subset, successful treatment was defined as a viral load below a detection limit of 50 copies/ml at week 12 after starting treatment. The additional data amounted to 19 patients for the DRV analysis and 112 for the quadruple NRTI analysis. The data sets were analyzed using the statistical software suite R [22].

Results

As an example of the influence of data generated by the bioinformatic interpretation approach geno2pheno_[resistance] on the HIV-GRADE rules, selected mutation scores are shown in table 1. While mutations like T215Y/F and K70R or an insertion at position 69 add up to higher scores and can therefore lead to a classification as an AZT-resistant virus strain, mutations M184I or M184V decrease the score and can lead, in case of mutations with low positive scores, to a classification as a fully AZT-susceptible virus strain. Although the mutation K65R shows no or only little influence on AZT susceptibility in the geno2pheno_[resistance] scores, a resensitizing effect is described in the literature [6]. These additional data and other published data were used to establish the HIV-GRADE rules for AZT (table 2).

Using the algorithm specification interface, the complete HIV-GRADE algorithm can be described using 400 instructions.

Other rules-based interpretation systems need a lower number of instructions to describe their rules. While HIVdb needs only 22 instructions, REGA needs 67 instructions (most of them are very complex). The ANRS rules can be described with 104 instructions, but this algorithm lacks the implementation of resensitizing effects.

Coefficients of determination (R^2) were calculated using R on the 334 TCEs of the basic validation data set. HIV-GRADE ($R^2 = 0.40$) (fig. 1) and Stanford HIVdb ($R^2 = 0.40$) showed the highest coefficients of determination, followed by REGA algorithm ($R^2 = 0.36$) and ANRS ($R^2 = 0.35$). On the much smaller data set for DRV coefficients of determination were 0.36 for HIV-GRADE, 0.33 for ANRS and 0.32 for Stanford HIVdb. REGA algorithm was not tested using this data set. The performance of HIV-GRADE was superior in predicting the quadruple NRTI treatment outcome. The accuracy of HIV-GRADE drug combination rules for predicting treatment outcome of the 112 patients in this special data set was 83% compared to HIV-GRADE single drug rules (70%) or other interpretation systems like REGA (71%), HIVdb (70%) and ANRS (69%).

Discussion

The HIV-GRADE rules-based interpretation system shows at least an equal performance to the other publicly available rules-based drug resistance interpretation systems on our basic clinical database. It shows an improved performance in special problems like quadruple NRTI treatment or in interpreting drug resistance-associated mutations in some of the more recently available drugs like DRV. While the quadruple NRTI treatment is a particularly good example for the prediction efficacy of a complete antiretroviral treatment, there is still room for improving the prediction by including information about the synergistic effects in treatment combinations [23]. At the moment, combination effects are only assessed for the NRTIs in the HIV-GRADE rule base, but expansion to all kinds of drug combinations is under investigation. Special combinations like the quadruple NRTI treatment or other special combinations like AZT and 3TC are analyzed as one treatment component, and the result of these analyses is provided in an extra section of the output. Each drug is analyzed independently with and without the effect of resensitization, and in the output of the HIV-GRADE tool there are 2 interpretations shown for the relevant drugs. The first is the interpretation of drug resistance-associated mutations without taking the resen-

Fig. 1. Graphical representation of the multiple linear regression model applied to the HIV-GRADE algorithm. The y-axis represents the logarithmic difference of the viral load before and after treatment. Thus, values above zero denote a viral load decline 12 weeks after treatment change. The x-axis shows the difference in activity between the current and previous treatments. This difference has been corrected by baseline viral load. ΔLOGVL = Difference in log viral load; $\Delta\text{ADS}_{\text{VLco}}$ = viral-load-corrected difference in active drug score.

sitizing effects into account, while the second integrates these effects. The latter interpretations are suffixed *_SP* in the output, which is an acronym for selective pressure and indicates that this interpretation may only be useful as long as the selective pressure on the resensitizing mutation is retained in a future regimen. The intention here is to remind that resistance-associated mutations can reverse to wild type, especially if the mutations lead to a disadvantage in viral replication [24]. Whereas the reappearance of wild-type mutations in cases of therapy interruption has been extensively reported [25, 26], the disappearance of resensitizing drug resistance mutations mediating higher susceptibility to a current regimen is not well described. The advantage of specific drug combinations, particularly one drug being resensitized by a mutation causing resistance to a certain second drug and this second drug only being administered to keep the selective pressure on this mutation, remains unclear. However, this assumption is supported by observations like the successful treatment of patients with the dual protease inhibitor combination-boosted lopinavir and saquinavir

[27], where the resensitizing effects of the mutation L76V seem to be relevant [28].

Although bioinformatic-based prediction tools like the EuResist engine [29, 30] have shown a superior performance when predicting the success of combinations of antiretroviral drugs, as compared to human experts [31] or rules-based algorithms [32], they are still not widely accepted. Rules-based systems or systems that predict the phenotypic resistance factor for each drug, like geno2pheno_[resistance] (virtual phenotype), are most frequently used. Also discussed by Lengauer [33] in this issue, one speculated reason for this ambiguous acceptance is that the mechanisms of how the bioinformatic systems choose the optimal treatment are sometimes not comprehensible. In many cases this could be prevented by an optimized presentation of the results. Even if some of the rules in the rules-based systems may not be comprehensible, they are at least verifiable in the case of a published rule set. Another reason might also be the conservative

selection of the best treatment option by tools like the EuResist engine or THEO [34]. Especially in cases where new drugs are considered to be a relevant option, they might be ignored by the bioinformatic systems because little or no TCEs with those drugs are available in the training database. Rules-based systems can circumvent these problems by open publication of the rule sets and the possibility of integrating current published data. HIV-GRADE tries to combine bioinformatic efforts and human experience, providing a tool that aims at giving clear and concise advice to clinicians. The parallel presentation of the results of several other relevant drug resistance interpretation systems provides additional safety to the users who can identify samples with uncertain interpretations when predictions among the systems are discordant. The HIV-GRADE tool is freely available on the Internet [14] and the XML file can be obtained by contacting the author.

References

- 1 Boyd MA, Hill AM: Clinical management of treatment-experienced, HIV/AIDS patients in the combination antiretroviral therapy era. *Pharmacoeconomics* 2010;28(suppl 1):17–34.
- 2 Preston BD, Poiesz BJ, Loeb LA: Fidelity of HIV-1 reverse transcriptase. *Science* 1988; 242:1168–1171.
- 3 Miller V, de Béthune MP, Kober A, Stürmer M, Hertogs K, Pauwels R, et al: Patterns of resistance and cross-resistance to human immunodeficiency virus type 1 reverse transcriptase inhibitors in patients treated with the nonnucleoside reverse transcriptase inhibitor loviride. *Antimicrob Agents Chemother* 1998;42:3123–3129.
- 4 De Luca A, Di Giambenedetto S, Maserati R, Gianotti N, Narciso P, Antinori A, et al: Interpretation of genotypic HIV-1 resistance to darunavir and virological response: validation of available systems and of a new score. *Antivir Ther (Lond.)* 2011;16:489–497.
- 5 Götte M, Wainberg MA: Biochemical mechanisms involved in overcoming HIV resistance to nucleoside inhibitors of reverse transcriptase. *Drug Resist Updat* 2000;3:30–38.
- 6 Wolf K, Walter H, Beerenwinkel N, Keulen W, Kaiser R, Hoffmann D, et al: Tenofovir resistance and resensitization. *Antimicrob Agents Chemother* 2003;47:3478–3484.
- 7 Tural C, Ruiz L, Holtzer C, Schapiro J, Viciana P, González J, et al: Clinical utility of HIV-1 genotyping and expert advice: the Havana trial. *AIDS* 2002;16:209–218.
- 8 Durant J, Clevenbergh P, Halfon P, Delguidice P, Porsin S, Simonet P, et al: Drug-resistance genotyping in HIV-1 therapy: the VIRADAPT randomised controlled trial. *Lancet* 1999;353:2195–2199.
- 9 Cingolani A, Antinori A, Rizzo MG, Murri R, Ammassari A, Baldini F, et al: Usefulness of monitoring HIV drug resistance and adherence in individuals failing highly active antiretroviral therapy: a randomized study (ARGENTA). *AIDS* 2002;16:369–379.
- 10 Meynard J-L, Vray M, Morand-Joubert L, Race E, Descamps D, Peytavin G, et al: Phenotypic or genotypic resistance testing for choosing antiretroviral therapy after treatment failure: a randomized trial. *AIDS* 2002; 16:727–736.
- 11 Vercauteren J, Vandamme A-M: Algorithms for the interpretation of HIV-1 genotypic drug resistance information. *Antivir Res* 2006;71:335–342.
- 12 Beerenwinkel N, Däumer M, Oette M, Korn K, Hoffmann D, Kaiser R, et al: Geno2pheno: estimating phenotypic drug resistance from HIV-1 genotypes. *Nucleic Acids Res* 2003;31: 3850–3855.
- 13 HIV Drug Resistance Database. <http://hivdb.stanford.edu/> (accessed July 5, 2011).
- 14 HIV-GRADE: Homepage. <http://www.hiv-grade.de/cms/grade/> (accessed July 5, 2011).
- 15 Huang X, Zhang J: Methods for comparing a DNA sequence with a protein sequence. *Comput Appl Biosci* 1996;12:497–506.
- 16 Betts BJ, Shafer RW: Algorithm specification interface for human immunodeficiency virus type 1 genotypic interpretation. *J Clin Microbiol* 2003;41:2792–2794.
- 17 HIV French Resistance – HIV-1 genotypic drug resistance interpretation's algorithms. <http://www.hivfrenchresistance.org/> (accessed July 5, 2011).
- 18 Lab: Rega HIV drug resistance algorithms. <http://jose.med.kuleuven.be/lab/index.php?id=30> (accessed July 5, 2011).
- 19 Geno2pheno resistance. <http://www.geno2pheno.org/> (accessed July 5, 2011).
- 20 Prosperi MC, Di Giambenedetto S, Fanti I, Meini G, Bruzzone B, Callegaro A, et al: A prognostic model for estimating the time to virologic failure in HIV-1 infected patients undergoing a new combination antiretroviral therapy regimen. *BMC Med Inform Decis Mak* 2011;11:40.
- 21 Zazzi M, Prosperi M, Vicenti I, Di Giambenedetto S, Callegaro A, Bruzzone B, et al: Rules-based HIV-1 genotypic resistance interpretation systems predict 8 week and 24 week virological antiretroviral treatment outcome and benefit from drug potency weighting. *J. Antimicrob Chemother* 2009; 64:616–624.
- 22 Team RDC. R: A language and environment for statistical computing. Vienna, 2009. <http://www.R-project.org>.

- 23 Feng JY, Ly JK, Myrick F, Goodman D, White KL, Svarovskaia ES, et al: The triple combination of tenofovir, emtricitabine and efavirenz shows synergistic anti-HIV-1 activity in vitro: a mechanism of action study. *Retrovirology* 2009;6:44.
- 24 Pinggen M, Nijhuis M, de Bruijn JA, Boucher CAB, Wensing AMJ: Evolutionary pathways of transmitted drug-resistant HIV-1. *J Antimicrob Chemother* 2011;66:1467–1480.
- 25 Verhofstede C, Wanzele FV, van der Gucht B, de Cabooter N, Plum J: Interruption of reverse transcriptase inhibitors or a switch from reverse transcriptase to protease inhibitors resulted in a fast reappearance of virus strains with a reverse transcriptase inhibitor-sensitive genotype. *AIDS* 1999;13:2541–2546.
- 26 Walter H, Löw P, Harrer T, Schmitt M, Schwingel E, Tschochner M, et al: No evidence for persistence of multidrug-resistant viral strains after a 7-month treatment interruption in an HIV-1-infected individual. *J Acquir Immune Defic Syndr* 2002;31:137–146.
- 27 Staszewski S, Babacan E, Stephan C, Haberl A, Carlebach A, Gute P, et al: The LOPSAQ study: 48-week analysis of a boosted double protease inhibitor regimen containing lopinavir/ritonavir plus saquinavir without additional antiretroviral therapy. *J Antimicrob Chemother* 2006;58:1024–1030.
- 28 Wiesmann F, Vachta J, Ehret R, Walter H, Kaiser R, Stürmer M, et al: The L76V mutation in HIV-1 protease is potentially associated with hypersusceptibility to protease inhibitors atazanavir and saquinavir: is there a clinical advantage? *AIDS Res Ther* 2011;8:7.
- 29 Altmann A, Rosen-Zvi M, Prosperi M, Aharoni E, Neuvirth H, Schülter E, et al: Comparison of classifier fusion methods for predicting response to anti HIV-1 therapy. *PLoS ONE* 2008;3:e3470.
- 30 EuResist Data Analysis. <http://engine.euresist.org/> (accessed July 5, 2011).
- 31 Zazzi M, Kaiser R, Sönnernborg A, Struck D, Altmann A, Prosperi M, et al: Prediction of response to antiretroviral therapy by human experts and by the EuResist data-driven expert system (the EVE study). *HIV Med* 2011;12:211–218.
- 32 Prosperi MCF, Altmann A, Rosen-Zvi M, Aharoni E, Borgulya G, Bazso F, et al: Investigation of expert rule bases, logistic regression, and non-linear machine learning techniques for predicting response to antiretroviral treatment. *Antivir Ther (Lond.)* 2009;14:433–442.
- 33 Lengauer T: Bioinformatical assistance of selecting anti-HIV therapies: where do we stand? *Intervirology* 2012;55:108–112.
- 34 Altmann A, Däumer M, Beerenwinkel N, Peres Y, Schülter E, Büch J, et al: Predicting the response to combination antiretroviral therapy: retrospective validation of geno2pheno-THEO on a large clinical database. *J Infect Dis* 2009;199:999–1006.