Discordant predictions of residual activity could impact dolutegravir prescription upon raltegravir failure

Kristof Theys a,∗, Ana Abecasis b, Pieter Libin a,c, Perpeitua Gomes d,e, Joaquim Cabanas d, Ricardo J. Camacho a, Kristel Van Laethem a,f

a KU Leuven - University of Leuven, Department Microbiology and Immunology, Rega Institute for Medical Research, B-3000 Leuven, Belgium
b Global Health and Tropical Medicine, International Public Health and Biostatistics, Institute for Hygiene and Tropical Medicine, Universidade NOVA de Lisboa, Lisbon, Portugal
c Artificial Intelligence Lab, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
d LMCBH, SFC, HEM - Centro Hospitalar de Lisboa Ocidental, Lisbon, Portugal
e Centro de Investigación Interdisciplinar Egan Moniz (CIEM), Instituto Superior de Ciências da Saúde Egan Moniz, Monte de Caparica, Portugal
f University Hospitals Leuven, AIDS Reference Laboratory, Herestraat 49, 3000 Leuven, Belgium

A R T I C L E   I N F O

Article history:
Received 19 March 2015
Received in revised form 3 July 2015
Accepted 25 July 2015

Keywords:
dolutegravir resistance
HIV-1 interpretation
integrase treatment

A B S T R A C T

Background: Dolutegravir is approved for the treatment of HIV-1 patients exposed to other integrase inhibitors, but the decision to use dolutegravir in this setting should be informed by drug resistance testing.

Objectives: This study determined the extent of disagreement in predicted residual dolutegravir activity after raltegravir use, and identified individual mutational patterns for which uncertainty exists among HIV-1 expert systems.

Study design: Mutation patterns were classified in raltegravir signature pathways including positions 143, 148 and 155, and interpreted into clinically informative resistance levels using genotypic drug resistance interpretation systems ANRS v24, HIVdb v7.0 and Rega v9.1, and instructions of dolutegravir use as approved by the Food and Drug Administration and the European Medicines Agency.

Results: In 216 HIV-1 patients failing raltegravir-therapy, 87% patients displayed mutations associated with resistance towards integrase inhibitors. A total of 141 unique mutational patterns were observed, with N155H (25.4%), Q148H (16.2%) and Y143R (8.3%) the most prevalent signature mutations. The Q148 pathway occurred almost exclusively in HIV-1 subtype B viruses. Concordances in predicted dolutegravir susceptibility scores among 5 systems were obtained in 57.8% of patients, and concordant intermediate resistant and concordant resistant scores were only observed in 6.5% and 0.9% of patients, respectively. However, systems individually scored higher levels of dolutegravir intermediate resistance and resistance, ranging from 4.2% to 10.2% and from 14.8% to 22.7% of patients, respectively. Consensus on interpreting the extent of residual activity was lacking in 34.7% of patients and was highly resistance pathway-specific.

Conclusions: Dolutegravir may potentially be effective in the majority of HIV-1 patients failing raltegravir, but concern over the uncertainty in predicted residual activity could withhold clinicians from prescribing dolutegravir during its clinical assessment.

© 2015 Elsevier B.V. All rights reserved.

1. Background:

In absence of a therapeutic cure or a preventive vaccine for infection with Human Immunodeficiency Virus type 1 (HIV-1), potent virological suppression by antiretroviral treatment (ART) remains a key mechanism to reduce HIV-1 morbidity and mortality as well as to prevent further epidemic spread. Therapeutic options for HIV-1 infection strongly enlarged with the introduction of integrase (IN) strand transfer inhibitors (INSTIs) into clinical practice. This latest approved drug class includes raltegravir, elvitegravir and dolutegravir, which are currently all part of recommended first-line regimens [1,2].

With the recent advent of dolutegravir, long-term management of HIV-1 infection further improved as this second-generation INSTI...
was approved for use in patients previously exposed to other INSTIs [3,4]. Whereas extensive cross-resistance precludes sequential use of raltegravir and elvitegravir, a clinical benefit of dolutegravir following first-generation INSTI-containing regimen failure is supported by observations of retained in vitro activity against some raltegravir- or elvitegravir-resistant isolates [5–8]. Furthermore, the VIKING-3 and VIKING-4 clinical trials assessed dolutegravir efficacy in HIV-1 patients with a history of genotypic evidence of multi-class resistance including INSTIs [9–11].

The decision to use dolutegravir in this setting should be informed by genotypic drug resistance testing of the IN coding region [1–4]. Based on genotypic and phenotypic studies, viral adaptation to INSTI selective pressure occurs along distinct but not mutually exclusive pathways, which are characterized by signature resistance-associated mutations (RAMs) evolving into complex patterns during prolonged drug exposure [12–20]. However, a wide spectrum of identified INSTI RAMs and only limited reports on their impact on in vitro resistance or virological response challenge the clinical interpretation of dolutegravir resistance and therefore can affect optimization strategies for sequential use of cART.

2. Study objectives:

The objective of this study was to determine the extent of disagreement in predicted residual dolutegravir activity in order to improve the interpretation of genotypic dolutegravir resistance. The identification of individual mutational patterns associated with uncertainty among HIV-1 expert systems can provide valuable information for the design of in vitro and in vivo outcome studies to evaluate specific expert rules of the current resistance interpretation algorithms.

3. Study design:

HIV-1 IN sequences from 216 raltegravir-experienced patients were retrieved from the Portuguese, Leuven University and Stanford University HIV-1 drug resistance databases [21,22]. The most recent sequence was retained when multiple isolates of a patient were available. Viral subtyping was performed by the Rega subtyping tool v3.0 [23].

INSTI RAMs were defined by their presence in the IAS-USA 2014 list [24], in drug resistance interpretation systems ANRS v24 [25], HIVdb v7.0.1 [26] and Rega v9.1.0 [27], or in FDA and EMA package inserts of the three INSTIs [3,4,28,29]. The following exhaustive mutation list was designed: S17N, H51Y, T66AIK, L68IV, L74IM, E92GQV, Q95KLR, Y143HR, H114Y, G118R, F121Y, E124A, E138ADKT, G140ACS, Y143AGHKKRS, P145S, Q146KLR, S147G, Q148EGHKRK, V151AI, S153AFY, M154I, N155HST, K156N, E157PQ, G163EKQRS, V165I, H183P, G193ER, T206S, V226CDEH, S230NR, D232N, R263K and V281 M, with T66AIK, E92GQV, G118R, F121Y, E138K, G140ACS, Y143AGHKKRS, P145S, Q146KLR, S147G, Q148EGHKRK, V151L, S153Y, N155HST, E157Q and R263K defined as major mutations against at least one of the three INSTI approved for clinical use. In particular, the following 40 mutations were listed to be associated with reduced susceptibility to dolutegravir: H51Y, T66K, L74IM, E92GQV, Q95L, Y143HR, H114Y, G118R, F121Y, T124A, E138ADKT G140ACS, Y143CHR, Q148HKRK, V151L, S153Y, M154I, N155H, E157PQ, G163EKQRS, G193ER, R263K. Polymorphic INSTI resistance-associated mutations were identified based on Rhee et al. [30].

Mutation patterns were categorized into 6 groups according to mutations at raltegravir signature positions 143, 148 and 155 [31], and assigned clinically informative resistance levels of susceptible (S), intermediate resistant (IR) or resistant (R), as scored by ANRS, HIVdb and Rega. For HIVdb, five level scores were simplified to three levels: susceptible and potential low-level resistant were scored as susceptible, low-level and intermediate resistant were scored as intermediate resistant and high-level resistant as resistant. As independent data of in vivo dolutegravir response data from outside clinical trials is still limited, two additional categorization schemes were derived from information provided in the FDA and EMA package inserts [3,4]. According to FDA, poor virologic response was observed in VIKING-3 study patients displaying Q148HR and at least 2 mutations of L74IM, E138ADKT, G140AOS, Y143HR, E157Q, G163EKQRS or G193ER, or in patients having at least 3 mutations of L74IM, E138ADKT, G140AOS, Y143HR, Q148HR, E157Q, G163EKQRS or G193ER (respectively 18% and 25% of patients obtained <50 RNA copies/ml at week 24), and therefore these mutational patterns were categorized as resistant in this study. A diminished response was observed in patients displaying Y143CHR without any Q148 mutation or displaying Q148HR and G140AOS without any other INSTI mutation (<50 RNA copies/ml at week 24 in 56% of patients), and these mutational patterns were categorized as intermediate resistant. All other mutational combinations were scored susceptible. In a parallel analysis for EMA, mutation patterns associated with reduced in vivo response rates were translated into resistant for any Q148 mutation and at least 2 mutations of L74I, E138AKT or G140ACS (25% of patients obtained <50 copies/ml at week 24), into intermediate resistant for any Q148 mutation and 1 mutation of L74I, E138AKT or G140ACS (65% of patients obtained <50 RNA copies/ml at week 24) and into susceptible for any other combination of mutations. The interpretation according to EMA package insert was consistent with patterns published in Castagna et al. [27,31].

A prediction for raltegravir was defined concordant among the three interpretation algorithms when they all scored identical resistance levels, and was defined discordant otherwise. A prediction for dolutegravir was defined concordant among the three interpretation algorithms and the two categorization schemes when all systems scored identical resistance levels, and was defined discordant otherwise. Alternatively, as for raltegravir, discordances in predicted dolutegravir activity were also examined using only the three interpretation algorithms. In addition to these percentage agreements, the Krippendorff's alpha reliability coefficient (α) was used to obtain chance-corrected measures of inter-rater agreement that accounted for the ordinal scaling of predictions by multiple raters, as implemented in the R statistical software package [32]. An agreement coefficient α between 0.67 and 1.00 was considered an indication for good agreement among systems.

4. Results:

Forty-three of the 79 defined INSTI RAMs (54%) were detected in 216 raltegravir-experienced patients (Figure 1). A total of 188 HIV-1 sequences (87%) carried at least one INSTI mutation and evidence of a major INSTI mutation was observed in 132 patients (61%). A signature mutation N155H displayed the highest prevalence (25.4%), followed by Q148H (16.2%) and Y143R (8.3%). Most prevalent major INSTI RAMs were G140S (18.5%), E157Q (6.9%) and E92Q (4.6%), of which G140S was mainly associated with Q148H (88%), while other INSTI mutations T124A (31.4%), S17N (25.2%), T206S (25.0%), Y143H (15.2%), V151I (11.6%) and K156N (9.7%) prevailed.

Algorithm-specific resistance levels for raltegravir were for ANRS (S=41.7%, R=58.3%), HIVdb (S=36.6%, I=7.9%, R=55.6%) and Rega (S=42.6%, I=5.1%, R=52.3%) (Figure 2A). Predicted genotypic susceptibility levels for raltegravir displayed 8.8% discordances between ANRS, HIVdb and Rega, while discordant predictions were
resistant (52.3%), intermediate resistant (2.3%) and susceptible (36.6%) (Figure 2C). System-specific resistance levels for dolutegravir (Figure 2B) were for ANRS (S=75.9%, I=14.8%, R=9.3%), HIVdb (S=69.0%, I=22.7%, R=8.3%), Rega (S=70.4%, I=19.4%, R=10.2%), FDA (S=69.0%, I=20.8%, R=10.2%) and EMA (S=79.6%, I=16.2%, R=4.2%). Discordances in predicted dolutegravir activity between the three interpretation algorithms and the two derived categorization schemes was observed in 34.7% of patients, while 57.8%, 6.5% and 0.9% of the sequences were concordantly scored susceptible, intermediate resistant and resistant, respectively (Figure 2C). The frequency of discordant predictions decreased to 23.1% of patients when only the three interpretation algorithms were considered, with α values of 0.72 and 0.81 for agreement between the five and three systems respectively (Figure 2C, Table 1).

A total of 141 unique INSTI mutational patterns were classified into 6 groups according to mutations at raltegravir signature positions 143, 148 and 155 (Figure 3). The Y143 group contained 23 patterns observed in 24 patients (11.1%), with mutations Y143R (70.8%), Y143C (33.3%), Y143H (12.5%) and Y143S (4.2%). These mutations were accompanied by T97A (75.0%), T206S (37.5%), L74M (37.5%), T124A (16.7%), G163R (16.7%) and S230R (16.7%). An S score was assigned in 69.2% of all individual predictions compared to IR in 25.8% and R in only 5.0%. The Q148 group contained 31 patterns observed in 40 patients (18.5%), with signature mutations Q148H (82.5%), Q148R (15%) and Q148K (2.5%). Prevailing mutations were G140S (85.0%), K156N (22.5%), T124A (17.5%) and T206S (12.5%). An S score was assigned in only 3.0% of all individual predictions compared to 68.0% and 29.0% for IR and R scores.
The N155H group contained 41 patterns in 47 patients (21.8%), with mutations T124A (46.8%), V151I (42.6%), T206S (23.4%), T97A (17.0%) and G163R (17.0%), and was characterized by a high frequency of S scores (89.4%). As a fourth group, 9 patients (4.2%) displayed distinct patterns that included multiple signature mutations. High-level resistance (R) by any algorithm was estimated in 35.6% of predictions compared to 31.1% for IR and 33.3% for S. A fifth group of 76 patients (35.2%) displayed 37 patterns without any signature mutation, which were predominantly scored S (97.6%). Of interest, 15.8% of patients showed other major INSTI RAMs such as T66A (n=1), T66I (n=1), E92Q (n=1), E138K (n=1), E157Q (n=5) and R263K (n=5). Finally, no INSTI mutation was observed in 20 patients (9.3%).

Disagreement in estimated activity of dolutegravir between the scoring systems was further evaluated by the extent of discordant predictions for each of the six patient groups (Figure 3 and Table 1). Patients displaying the N155H pathway showed a high level of concordant susceptible scores (66%) but a consensus was not detected for the remaining 34% of patients. High levels of discordant predictions were observed for patients displaying the Y143 pathway (95.8%), the Q148 pathway (62.5%) and multiple signature mutations (88.9%), resulting in low levels of concordantly estimated susceptible scores. The Q148 pathway was concordantly estimated intermediate resistant or resistant in 35.0% and 2.5% of patients respectively. The patient group with no signature INSTI RAMs showed a high percentage agreement in susceptible scores by all algorithms (96.1%). For 14 patients (6.5%), all three levels of predicted drug susceptibility (S, IR and R) were assigned, with in particular the Y143 patient group (16.7%), the N155H group (8.5%) and patients with multiple signature mutations (22.2%) most affected. When assessing discords only among the three expert-based interpretation algorithms, the percentage of discordant predictions decreased strongly for the Y143 pathway (41.7%) and patients with multiple signature mutations (33.3%), and to a lesser extent for the Q148 pathway (45.0%) (Table 1). No changes were observed for the N155H group and patients without any signature INSTI mutations. High values of chance-corrected agreement measures were only observed for the patient group with multiple signature mutations (\(\alpha = 0.88\)) and for the patient group that lacked a signature INSTI mutation (\(\alpha = 0.99\)), when limited to the three algorithms (Table 1).

Subtype B viruses were detected in 71% of patients. Predicted genotypic susceptibility levels of raltegravir were comparable between B and non-B infected patients with respect to concordant
Fig. 3. The frequency of concordant and discordant dolutegravir resistance scores in HIV-1 sequences obtained after raltegravir failure. Interpretation of drug activity was performed with rules derived from ANRS v24, HIVdb v7.0.1, Rega v9.1.0, FDA and EMA package inserts, as defined in text. The number of unique mutational patterns are displayed and categorized according to the presence of signature mutations Y143CHR, Q148H/K/R and/or N155H. Concordant susceptible □, intermediate resistant □, resistant □ and discordant □ scores.
susceptible scores (35.3% vs 39.7%, p-value = 0.65) and discordant scores (8.5% vs 9.5%, p-value = 1). Proportions of predicted genotypic susceptibility levels for dolutegravir varied between B versus non-B with respect to discordant susceptible scores (51.6% vs 73.0%, p-value = 0.006), while discordant scores were not differently prevalent (37.9% vs 27.0%, p-value = 0.16) for all five systems. Similar findings were obtained for the three expert-based algorithms only, with respect to discordant susceptible (56.9% vs 84.1%, p-value = 0.002) and discordant scores (26.1% vs 15.9%, p-value = 0.15). Patterns within the Q148 group were predominantly found in HIV-1 subtype B infected patients (95%) compared to N155H (58%) and Y143 (67%) groups.

5. Discussion:

Additionally to more potent and tolerable HIV-1 inhibitors that have gradually entered the market in last decades, HIV-1 inhibitors that retain activity against drug resistant HIV-1 variants remain a keystone in the life-long treatment of HIV-1 patients. Due to therapeutic failure, the treatment of HIV-1 infected patients is generally characterized by multiple consecutive drug regimens. The emergence of drug-resistant HIV-1 variants impairs the effectiveness of similar HIV-1 inhibitors and complicates treatment sequencing strategies as resistance profiles can be shared between inhibitors from the same drug class [33]. In clinical practice, the management of HIV-1 drug resistance constitutes an essential part of long-term successful treatment of HIV-1 infection. However, the translation of genotypic drug resistance into clinical implications is a challenging task, and drug resistance interpretation systems are widely used to assist clinicians in the management of individual patients for whom treatment switches are indicated. Uncertainty in predicted dolutegravir activity could have implications for genotype-guided treatment-decisions. Hence, information on the mutational patterns underlying discordances in estimated outcome serves as a valuable prior for directed phenotypic resistance experiments and for enabling experts to refine their interpretation systems.

Estimates of dolutegravir activity were obtained using three genotypic drug resistance interpretation algorithms, widely used in the clinical management of HIV-1 infected patients. Additionally, two categorization schemes were derived from the company package inserts, given that dolutegravir response data outside the VIKING clinical trials is not abundantly available to date. Based upon our study and others, dolutegravir may be potentially effective in the majority of HIV-1 patients failing a raltegravir–containing regimen [6,7,9,15]. However, a consensus on interpreting the extent of residual activity among the five systems was lacking in a large proportion of patients and showed a resistance pathway-specific pattern, which could lead to uncertainty in individual patient management. Although discordant intermediate resistance and resistance scores were only obtained in 7.4% of all patients, individual interpretation systems scored these resistance levels more often, from 20.4% for the FDA scheme to 31% for HIVdb and FDA scheme. The highest levels of disagreement were observed for Y143 and Q148 groups and for patterns including multiple signature mutations. The Y143 group showed susceptible scores for almost 70% of all predictions, but discordant susceptibility scores were only obtained for 4% of predictions. The intermediate resistance score that was triggered by the FDA scheme when Y143CHR was present in absence of any Q148 mutation mainly caused the high level of discordance in this group. The Q148 group was predicted to be most affected, with only 3% of all individual predictions assigning a susceptible score, in line with reports that the largest impact on viral susceptibility and virological response in clinical trials was observed for study patients with Q148 mutations in association with other INSTIs resistance-associated mutations (RAMs) [9,10]. This group also showed a high percentage of discordance between systems. The N155H pathway showed a high proportion of individual S scores and lower discordance rates. Similar findings were also observed for patients that did not display any of the raltegravir signature mutations. However, 3 patients within this group displayed the R263K mutation, which has been reported to be selected in vitro and in vivo by both raltegravir and dolutegravir [17,34]. This non-polymorphic mutation is scored high-level resistant by ANRS while intermediate resistant by HIVdb and Rega, being discordantly estimated in all patients. Finally, the presence of multiple signature mutations were associated with a low prevalence of predicted S scores and a high level of discordant estimates, confirming that more complex mutational patterns are often associated with increased challenges for clinical interpretation [33].

When the analysis was limited to the three interpretation algorithms only, in order to account for possible distortions resulting from the inclusion of the two categorization schemes, the frequency
of discordant predictions decreased from 34.7% to 23.1%, while concordant predictions were more observed for all three resistance scores. The finding that raltegravir activity was discordantly predicted in merely 8.8% of patients could highlight the lack of sufficient knowledge on dolutegravir resistance, and supports the study rationale of including the package insert-derived categorization schemes. The Y143, Q148 and multiple signature mutation patient groups showed lower levels of discordant predictions, although remaining above 40% for the Y143 and Q148 groups.

The extent of discordant predictions between the systems was examined using two different measures, the percentage of agreement and the Krippendorff’s alpha. The latter coefficient corrects for agreement occurring by chance and accounts for ordinal resistance scores. The highest alpha values were obtained for the patient group with multiple signature mutations and for the patient group without any signature mutation. Differences in agreement between the two measures can be expected, given that Krippendorff’s alpha calculates disagreements among raters instead of the percentage of agreement and applies a different weighting for more distant scores (susceptible versus resistant scores) than for less distant scores. Conclusions on agreement should however be carefully made, since guidelines for the interpretation of reliability coefficients vary and a consensus on the agreement statistic to be used as well as an acceptable level of agreement are not generally defined.

Several INSTI RAMs were highly prevalent in our study population but they represented consensus residues in specific subtypes (S17N, T124A, T203S) or known polymorphisms in treatment-naïve patients (T97A, V151I and K156S) [22]. The predominant presence of subtype B viruses in the Q148 pathway has been reported before [15], and this preference is most likely explained by a lower genetic barrier for INSTI RAMs at position 140, which were almost exclusively found in combination with raltegravir signature mutations Q148. Because of a different codon usage, subtype B viruses require only a single nucleotide substitution for the change G140S while two substitutions are required in non-B subtypes [35]. Given that this pathway was associated with high levels of R predictions but also disagreement, this could have implications for dolutegravir use.

A comparison of predictions by the different systems represented an essential aspect of this study. However, the predictive performance of each system was not addressed in this study since virological endpoint references were not available, and the clinical implications of our findings should be further investigated. Our study used three well-acknowledged genotypic resistance interpretation algorithms, but other available HIV-1 interpretation systems should be investigated as well to fully characterize flaws in the translation of observed resistance to clinically informed treatment decisions. All patients in our study were experienced with other drug classes before raltegravir treatment, reflecting that raltegravir was initially limited to therapy-experienced patients with multi-class drug resistance against reverse transcriptase and protease inhibitors. To date, raltegravir is also recommended for first-line treatment and this might lead to other frequencies of INSTI mutations, because the activity of the entire regimen is higher in this setting and clinicians are less permissive in maintaining sub-optimal therapeutic pressure once virological failure is detected. Consequently, this will affect the residual activity of dolutegravir upon raltegravir failure.

In conclusion, despite extensive reports on INSTI resistance development, our findings strongly indicate that further research is necessary for a better understanding of genotypic resistance towards dolutegravir, in particular for mutation pathways involving positions 143 and 148. Besides collecting and analyzing cohort data, open access to raw data from large clinical trials could enable independent re-analysis and might eventually improve personalized treatment of HIV-1 patients.

Conflict of interest
None.

Funding
This work was supported by the Fonds voor Wetenschappelijk Onderzoek Vlaanderen (G.0692.14N) and by the AIDS Reference Laboratory of Leuven that receives support from the Belgian Ministry of Social Affairs through a fund within the Health Insurance System. KT is funded as a Postdoctoral Researcher by the Research Foundation – Flanders (FWO). The computational resources and software used in this work were provided by the Hercules Foundation and the Flemish Government – department EWI-FWO Krediet aan Novorsers (Thys, KAN2012 1.5.249.12.). This work was partly supported by BEST HOPE: Bio-Molecular and Epidemiological Surveillance of HIV Transmitted Drug Resistance, Hepatitis Co-Infections and Ongoing Transmission Patterns in Europe (project funded through HIVERA: Harmonizing Integrating Vitalizing European Research on HIV/Aids, grant 249,697) and by L’Oréal Portugal Medals of Honor for Women in Science 2012 (financed through L’Oréal Portugal, Comissão Nacional da Unesco and Fundação para a Ciência e Tecnologia).

Ethical approval
The study protocol was conducted in accordance with the Declaration of Helsinki and approved by the ethical committees of UZ Leuven (B322201420270/S56109) and Centro Hospitalar de Lisboa Ocidental (108/CES-2014).

Acknowledgements
KT and KVL designed the study. KT, AB and PL performed the analysis. PG, JC, RJC and KVL collected the data. KT, AB, RJC and KVL wrote the manuscript. All authors contributed to the interpretation and discussion of the results. We thank all patients included in the study. We acknowledge the Portuguese HIV-1 Resistance Study Group and all personnel at the AIDS Reference Laboratory and Center in Leuven.

References


