Background: Nucleoside (or nucleotide) reverse transcriptase inhibitors (NRTIs) cause side effects in some patients, prompting the use of either partly or fully NRTI-sparing regimens.

Methods: We used data from the Swiss HIV Cohort Study to estimate the effectiveness of two new dolutegravir dual regimens relative to the alternative NRTI-sparing dual regimens that our clinicians used previously. We emulated two trials by propensity score matching case patients on the dolutegravir regimen with control patients on an alternative regimen. We analysed the case control sets using a Bayesian Cox model and estimated effectiveness as the percentage still on their trial regimen without virological failure at 48 weeks.

Results: In a comparison of partly NRTI-sparing regimens, 58 cases treated with dolutegravir were matched to 17 controls treated with boosted darunavir (both with lamivudine or emtricitabine). The estimated difference in effectiveness was 15% (95% credible interval [CrI] 2–33) and 12% (95% CrI 0–26) in two sequential analyses 1 year apart. In a comparison of fully NRTI-sparing regimens, 54 cases treated with dolutegravir were matched to 32 controls treated with raltegravir (both with boosted darunavir). The estimated difference in effectiveness was 9% (95% CrI 1–21) and 5% (95% CrI 4–15) in the two sequential analyses.

Conclusions: Estimates of relative effectiveness suggest that both dolutegravir regimens are not inferior to these alternative regimens. All four regimens seem suitable for patients needing an NRTI-sparing regimen: there were few virological failures and few treatment changes due to toxicity.

Introduction

Nucleoside (or nucleotide) reverse transcriptase inhibitors (NRTIs) are the backbone of combination antiretroviral therapy. But for some patients, cumulative exposure to drugs in this class has led to serious side effects [1,2] prompting the use of NRTI-sparing regimens. These patients have usually been given dual regimens that are either partly NRTI-sparing, with a single drug from this class (rather than two) plus an anchor drug, or fully NRTI-sparing with two anchor drugs.

Early dual regimens had higher rates of virological failure than three drug regimens [3–5]. Hence,
the combination of two NRTIs plus an anchor drug is still recommended for both an initial regimen and whenever changes to treatment are made [6]. Two regimens have been recommended for patients unable to take NRTIs – boosted darunavir with raltegravir and boosted lopinavir with lamivudine [6]. Other regimens are being investigated but are not yet recommended because of limited evidence. Different NRTI-sparing dual regimens have never been compared with each other in randomized trials [7,8] and indeed this may never happen in patients with side effects because of the difficulty of recruiting suitable patients. Because dual regimens are seldom used, there are few data from clinical practice with which to compare one dual regimen to another.

The efficacy, safety and resistance profile of dolutegravir has led to its recent use in NRTI-sparing dual regimens [9]. There is no evidence that dolutegravir-based dual regimens are better than the alternatives and in some jurisdictions the alternatives may be cheaper [10]. So there is value in using the observational data we have to provide clinicians with the best information currently available about the relative merits of alternative NRTI-sparing dual regimens.

In this study, we use data from the Swiss HIV Cohort Study (SHCS) to estimate the effectiveness of two new dolutegravir-based dual regimens (dolutegravir with lamivudine or emtricitabine; dolutegravir with boosted darunavir) relative to alternative NRTI-sparing dual regimens that clinicians used previously (boosted darunavir with lamivudine or emtricitabine; raltegravir with boosted darunavir). We use these data to emulate the randomized clinical trial that clinicians would prefer to see but which may never happen. We do not expect our results to provide definitive answers immediately: rather we show how we can periodically update results, re-estimating relative effectiveness as new data accrue.

Methods

Partly and fully NRTI-sparing regimens

We emulate two randomized trials in a switch setting, one between two partly NRTI-sparing regimens and the other between two fully NRTI-sparing regimens. In each emulated trial, we compare a dolutegravir-based dual regimen with its most frequently used earlier equivalent. The two partly NRTI-sparing regimens compared are dolutegravir and ritonavir-boosted darunavir where these anchor drugs were combined with either lamivudine or emtricitabine. The two fully NRTI-sparing regimens compared are dolutegravir and raltegravir where these integrase inhibitors were combined with boosted darunavir. With limited data for our emulated trials, we take a Bayesian approach so we can include additional information beyond our cohort’s data.

The SHCS is a prospective cohort with continuing enrolment of HIV-infected adults [11]. Follow-up visits are scheduled every 6 months with more frequent treatment monitoring (median 3.3 months between HIV RNA measurements). Patients from the SHCS were included in our emulated switch trials if they had HIV RNA below 50 copies/ml during the 6 months prior to starting a trial regimen (with one blip of less than 200 copies/ml allowed). Patients were excluded if they had documented resistance (HIVdb version 8.4 [12]) to any trial drug or if they had no resistance test but had in the past experienced virological failure on a regimen that included either a trial drug or any integrase inhibitor. Patients were also excluded if they started a trial regimen with a CD4+ T-cell count of below 200 cells/mm³, an AIDS defining disease in the previous year, chronic hepatitis B (hepatitis B surface antigen [HBsAg]-positive) or a past clinical event indicating severe hepatic impairment. There is some evidence that NRTI-sparing two drug regimens might be inferior to three drug regimens in patients with low CD4+ T-cell counts [13,14].

We evaluated the effectiveness of these regimens using the FDA’s guidance for industry [15] as a template. Patients were deemed to have failed on a regimen if HIV RNA was detected (a single measure above 1,000 copies/ml or the first of two consecutive measures greater than 50 copies/ml), if they discontinued any trial drug because of an adverse event or death, or if they discontinued any trial drug for other reasons. As in the FDA’s guidance, we specified a priori certain changes to the regimen that would not constitute failure: a switch from the non-dolutegravir trial regimen to the dolutegravir trial regimen; a switch from either trial regimen to a regimen that included tenofovir alafenamide, provided the new regimen had a lower pill count and there was evidence that the patient had been actively avoiding tenofovir disoproxil fumarate. Formal withdrawal from the SHCS was also not considered a failure.

Treatment assignment modelling

In a clinical trial, randomization assigns a known probability of treatment to each trial patient. With observational data, the probability of treatment is unknown and must be estimated. We estimated this probability – the propensity score – without reference to outcome data [16] using a logistic regression model with baseline covariates thought to represent clinical decision making. That is, there will be reasons why patients were given one of the two trial regimens and not the other, and the treatment assignment model must reflect those reasons.

In the partly NRTI-sparing trial, we considered that patients taking many additional medications, with a high Framingham risk score [17], with hepatitis C or failing previously on protease inhibitor regimens would be more likely to receive dolutegravir; while patients...
with risk factors for virological failure or previously failing on lamivudine or emtricitabine would be more likely to receive boosted darunavir. In the fully NRTI-sparing trial, we considered that patients taking many additional medications or with risk factors for virological failure would be more likely to receive dolutegravir; while patients with diabetes or depression would be more likely to receive raltegravir. In Additional file 1 (sections 3.1 and 3.2), we define each baseline covariate, explain our reasoning for its inclusion in the model and as a consequence of those reasons, the prior distribution we then assigned to each.

We used SAS/STAT 13.1 (SAS Institute Inc., Cary, NC, USA) for this modelling, using a BAYES statement in PROC GENMOD to fit a Bayesian logistic regression model.

Conceptual trial design and analysis
Provided the treatment assignment model is correctly specified, propensity scores can be used to design a conceptual trial such that for the patients included in the trial, the way treatments were assigned can be ignored in the analysis (as in a real trial). To create this conceptual trial, we matched by propensity score: for each patient starting the dolutegravir regimen (a ‘case’), we selected the patient starting the alternative regimen with the closest propensity score (a ‘control’), provided the control’s score was within a calliper. The calliper width was set to 0.2 standard deviations of the logit propensity score; a width recommended because of its theoretical properties and good performance in simulations [18]. This conceptual trial provides an answer to the question ‘what would happen if instead of treating this patient with the dolutegravir regimen, the alternative regimen was used instead?’ [19].

We matched patients in SAS using the PSMatching macro [20]. We expected cases to accrue faster than controls; therefore, we matched cases to controls with replacement. Re-using controls potentially allows closer matching between case and control patients and avoids new case patients being discarded because an otherwise suitable control had already been assigned to another case. The data for analysis were then a sample of case control sets, each set with one or more cases for each control. We analysed these sets using WinBUGs 1.4.3 to fit a Bayesian Cox model with a random effect for each control set (see the LeukFr example [21]). Our outcome was time to regimen failure, with follow-up censored when a regimen was changed for reasons that did not constitute failure or at the last follow-up visit (administrative censoring). We then used this model to estimate the difference in the percentage still on their trial regimen without virological failure at 48 weeks, the surrogate outcome used by the FDA for approval of antiretrovirals [15].

We add two sources of prior information to our analyses, both weakly informative [22]. First we placed an ‘uncertain direction’ prior – hazard ratio (HR) 1.0, 95% credible interval (CrI) 0.25–4.0 – on the estimate of relative effectiveness. This expressed our expectation that any difference in effectiveness between the two regimens would not be dramatic but within this interval. Second, we reviewed studies where patients switched onto the alternative regimens (boosted darunavir with lamivudine or emtricitabine [23–26], raltegravir with boosted darunavir [27–29]). Based on these studies, we expected a failure rate of approximately 10 failures per 100 patient-years on each control regimen. In our Bayesian Cox model, the baseline hazard function was represented by a gamma distribution. We discounted these historical data by a factor of 0.1 [30], to make our prior only weakly informative, setting a gamma prior equivalent to one failure in 10 patient-years. We did not use a formal meta-analysis to set this prior [31]; doing so might have been more than weakly informative.

We make two estimates for each emulated trial, the second a year after the first, to illustrate the increase in information with time. Carrying out an updated analysis is equivalent to a sequential trial. From a Bayesian perspective, there is no need to adjust credible intervals when making repeated estimates – the posterior estimate from an earlier analysis becomes the new prior for an updated analysis (equivalently all data from the emulated trial can be added to the original prior [32]).

Sensitivity analyses
The results of our original treatment assignment modelling suggested we had not adequately modelled the use of dolutegravir with lamivudine or emtricitabine. Clinicians may have used this regimen to avoid future toxicities, rather than in response to existing toxicities, and we did not anticipate this. We then fit an alternative treatment assignment model for the partly NRTI-sparing trial more consistent with this alternative treatment scenario (Additional file 1 section 3.3). We also re-fit our original treatment assignment model to two different patient subsets, where we tried to include only patients with existing toxicities (Additional file 1 section 3.4).

Results
Patients meeting study criteria
By June 2017, 122 and 157 patients had started a partly or fully NRTI-sparing regimen of interest in the SHCS (Figure 1). Of these, only 89 (73%) and 119 (76%), respectively, met study criteria. The most common reason for exclusion was detectable HIV RNA in the past 6 months (42% and 84% of those excluded on partly and fully NRTI-sparing regimes, respectively).
Of the patients meeting study criteria, those starting dolutegravir regimens tended to be younger and less likely to be taking co-medications associated with cardiovascular disease (Table 1). Differences in patient characteristics were greater between the two partly NRTI-sparing regimens: all comorbidities were more likely in patients starting the boosted darunavir regimen and those starting boosted darunavir had on average far more exposure to protease inhibitors and far less exposure to non-nucleoside reverse transcriptase inhibitors (9.7 and 0.6 years, respectively) than those starting dolutegravir (3.1 and 5.2 years, respectively).

Conceptual trials
In treatment assignment modelling of the two partly NRTI-sparing regimens, the data for some covariates were in conflict with the priors we asserted (Table A1 in Additional file 1). Posterior intervals were not contained within our wide prior intervals for three covariates (hepatitis C coinfection, more than three daily co-medications and previous failure on a protease inhibitor). The distribution of propensity scores from this model (Figure 2, left) shows a group of patients on darunavir with a low probability of receiving dolutegravir; these were of little use as controls because few patients receiving dolutegravir had similar scores. The conceptual trial then comprised 38 cases matched to 17 (44%) of the 29 available controls (Figure 1). For all covariates, case and control means were closer in the matched data than in the unmatched data, but some differences were still relatively large (Table A2 in Additional file 1). Data in conflict with our prior opinion...
and larger than expected differences between some covariate means for case and control patients led us to consider an alternative treatment assignment model in a sensitivity analysis.

In contrast, we were more confident of our treatment assignment modelling of the two fully NRTI-sparing regimens. The data were consistent with the priors we asserted (Table A3 in Additional file 1), propensity score distributions were similar for both regimens (Figure 2, right) and there were only trivial differences between case and control patient means for all covariates in the matched data (Table A4 in Additional file 1). The conceptual trial comprised 54 cases matched to 32 (49%) of the 65 available controls (Figure 1).

Initial and updated analyses
In both conceptual trials, the initial analysis suggested that patients starting dolutegravir regimens were more likely than patients starting the alternative regimens to still be on that regimen 48 weeks later with no evidence of virological failure (Table 2). At 48 weeks, an estimated 91% of patients would still be on dolutegravir with lamivudine or emtricitabine and an estimated 90% would still be on dolutegravir with boosted darunavir.

A year later, the number of patients starting dolutegravir with lamivudine or emtricitabine had increased from 77 to 124, with far more modest increases in the number of patients starting the other regimens (Figure 1). Matching led to many more cases in the two conceptual trials, but with little increase in the number of controls (from 17 to 22 and from 32 to 38 in the partly and fully NRTI-sparing trials, respectively). In both trials, the updated analysis gave attenuated estimates of effectiveness: at 48 weeks, an estimated 87% of patients would still be on dolutegravir with lamivudine or emtricitabine and an estimated 88% would still be on dolutegravir with boosted darunavir (Table 2).

### Table 1. Characteristics of patients (mean or percent) meeting inclusion criteria when starting a conceptual trial regimen using the data available for the initial analysis

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Partially NRTI-sparing</th>
<th>Fully NRTI-sparing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DTG + 3TC/FTC (n=60)</td>
<td>DRV/r + 3TC/FTC (n=29)</td>
</tr>
<tr>
<td>Demographic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, years</td>
<td>52</td>
<td>55</td>
</tr>
<tr>
<td>Female, %</td>
<td>27</td>
<td>34</td>
</tr>
<tr>
<td>Disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD4+ T-cell count, cells/µl</td>
<td>660</td>
<td>630</td>
</tr>
<tr>
<td>CD4+ T-cell nadir, cells/µl</td>
<td>230</td>
<td>230</td>
</tr>
<tr>
<td>Past AIDS diagnosis, %</td>
<td>20</td>
<td>41</td>
</tr>
<tr>
<td>Cumulative exposure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Since ART-naive, years</td>
<td>9.6</td>
<td>12</td>
</tr>
<tr>
<td>NNRTIs, years</td>
<td>5.2</td>
<td>0.6</td>
</tr>
<tr>
<td>PIs, years</td>
<td>3.1</td>
<td>9.7</td>
</tr>
<tr>
<td>Abacavir, years</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>Tenofovir, years</td>
<td>5.1</td>
<td>4.1</td>
</tr>
<tr>
<td>Comorbidities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hepatitis C, %</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>Depression, %a</td>
<td>20</td>
<td>32</td>
</tr>
<tr>
<td>Diabetes mellitus, %b</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>Past CVD event, %c</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>Framingham risk, %d</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Osteoporosis, %e</td>
<td>13</td>
<td>21</td>
</tr>
<tr>
<td>eGFR, ml/min/1.73 m²</td>
<td>90</td>
<td>76</td>
</tr>
<tr>
<td>Co-medications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drugs taken daily, n</td>
<td>1.7</td>
<td>2.3</td>
</tr>
<tr>
<td>Antiplatelet, %</td>
<td>15</td>
<td>24</td>
</tr>
<tr>
<td>Antihypertensive, %</td>
<td>25</td>
<td>55</td>
</tr>
<tr>
<td>Lipid lowering, %</td>
<td>23</td>
<td>38</td>
</tr>
</tbody>
</table>

*a* Any past patient report of depression or use of antidepressants. *b* Any past clinical diagnosis of diabetes, use of insulin or anti-diabetic medication, or a casual plasma glucose measured above 11.1 mmol/L. *c* Myocardial infarction, coronary angioplasty or stenting, coronary artery by-pass grafting, cerebral infarction (stroke), deep vein thrombosis, carotid endarterectomy, procedures on other arteries. *d* 10-year risk of coronary heart disease estimated using the Framingham risk score. *e* Any past low impact fracture or clinical diagnosis of osteoporosis. ART, antiretroviral therapy; CVD, cardiovascular disease; DRV/r, darunavir boosted with ritonavir; DTG, dolutegravir; eGFR, estimated glomerular filtration rate calculated using the CKD-EPI equation [45]; NNRTIs, non-nucleoside reverse transcriptase inhibitors; NRTI, nucleoside (or nucleotide) reverse transcriptase inhibitor; PIs, protease inhibitors; RGV, raltegravir; 3TC/FTC, lamivudine or emtricitabine.
After a year, the event rate in the raw data had increased among cases and decreased among controls. The update increased precision in both conceptual trials. The estimated HR had better precision (a lower standard error for the log HR; Table 2) but this led to only a modest decrease in the width of credible intervals for the difference between regimens at 48 weeks: from 31 to 26 percentage points in the partly NRTI-sparing trial, and from 22 to 19 percentage points in the fully NRTI-sparing trial.

In both trials, and in both analyses, credible intervals suggest that effectiveness on a dolutegravir regimen was no worse than 4 percentage points fewer at 48 weeks than on the alternative regimen.

Outcome composition
There were three deaths in our conceptual trials, five patients experienced virological failure and relatively few changes in regimen because of toxicity (Table 3). The three deaths on dolutegravir were due to a recurrent myocardial infarction, a cholangiocarcinoma and a pneumocystis pneumonia, respectively. Toxicities prompting a change to the regimen in more than a single patient were: dyslipidaemia (one patient on boosted darunavir with lamivudine or emtricitabine, two patients on dolutegravir with boosted darunavir), predominately nephrotoxicity (two patients on boosted darunavir with lamivudine or emtricitabine, one patient on dolutegravir with boosted darunavir), and...
predominately neurotoxicity (two patients on dolutegravir with lamivudine or emtricitabine, one patient on dolutegravir with boosted darunavir). The neurotoxicities for these three patients were described as: headache, neuropsychiatric and nervous system unspecified. Changes to trial regimens were consistent with the cautious use of dual regimens. Most changes were made with no evidence of virological failure or toxicity (Table 3).

Changes to trial regimens were consistent with the cautious use of dual regimens. Most changes were made with no evidence of virological failure or toxicity (Table 3). Among patients simplifying a trial regimen, 4 out of 7 (57%) and 20 out of 31 (65%) returned from partly and fully NRTI-sparing dual regimens, respectively, to three drug regimens that included tenofovir alafenamide.

Sensitivity analyses

Under an alternative treatment assignment model for the two partly NRTI-sparing regimens (Additional file 1 section 3.3), data for most covariates were consistent with the priors asserted (Table A5 in Additional file 1). The two regimens had again distinctly different propensity score distributions (Figure A1 in Additional file 1): a group of patients on dolutegravir had very high probabilities of receiving that regimen; few patients on darunavir had propensity scores that high. As a result, in the initial analysis, nine case patients were not matched with a control – eight of these had high propensity scores.

Table 2. Analyses of the two conceptual trials – both the initial analysis and the updated analysis 1 year later

<table>
<thead>
<tr>
<th>Measure</th>
<th>Partly NRTI-sparing regimens: DRV/r or DTG (both with 3TC/FTC)</th>
<th>Fully NRTI-sparing regimens: RGV or DTG (both with DRV/r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>On treatment and undetectable at 48 weeks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case regimen (with DTG), %</td>
<td>91 (81–97)</td>
<td>90 (81–96)</td>
</tr>
<tr>
<td>Control regimen, %</td>
<td>75 (56–88)</td>
<td>81 (68–90)</td>
</tr>
<tr>
<td>Difference (case - control), %</td>
<td>15 (2–33)</td>
<td>9 (-1–21)</td>
</tr>
<tr>
<td>Hazard ratio</td>
<td>0.34 (0.13–0.83)</td>
<td>0.51 (0.23–1.1)</td>
</tr>
</tbody>
</table>

Table 3. Outcomes for patients included in conceptual trials using the data available for the updated analysis

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Partly NRTI-sparing regimens</th>
<th>Fully NRTI-sparing regimens</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTG + 3TC/FTC (n=86)</td>
<td>DRV/r + 3TC/FTC (n=22)</td>
<td>DTG + DRV/r (n=71)</td>
</tr>
<tr>
<td>Censored</td>
<td>68</td>
<td>7</td>
</tr>
<tr>
<td>Administrative</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Simplification</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Failure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Death</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Virological failure</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Stopped – toxicity</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Stopped – other</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>New drugs added</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Data in parentheses are 95% credible intervals. *Dolutegravir with lamivudine or emtricitabine (partly nucleoside [or nucleotide] reverse transcriptase inhibitor [NRTI]- sparing); dolutegravir with boosted darunavir (fully NRTI-sparing). †Boosted darunavir with lamivudine or emtricitabine (partly NRTI-sparing); raltegravir with boosted darunavir (fully NRTI-sparing). DRV/r, darunavir boosted with ritonavir; DTG, dolutegravir; PY, patient years; RGV, raltegravir; 3TC/FTC, lamivudine or emtricitabine.

Antiviral Therapy

Page numbers not for citation purposes
scores. In the matched data, some difference between case and control means were still relatively large (Table A6 in Additional file 1). With more unmatched cases, the conceptual trial then had fewer cases than before but despite these differences, analyses led to almost identical estimates (Table A7 in Additional file 1).

Efforts to improve balance in the partly NRTI-sparing trial by restricting the trial to patients with existing toxicities were largely unsuccessful (Additional file 1 section 3.4). Additional restrictions reduced the number of events so that it was then difficult to assess either differences between data and priors or the balance between selected cases and controls. Estimates from two sensitivity analyses in different patient subsets (Table A8 in Additional file 1) were consistent with estimates made without the additional restrictions but were not of sufficient precision to really confirm those unrestricted estimates.

Discussion

Our results provide some evidence that these two dolutegravir regimens are not inferior to the alternatives previously used in the SHCS for patients needing an NRTI-sparing dual regimen. Our credible intervals suggest that at 48 weeks both dolutegravir regimens would be at worst 4% less effective than the alternatives – the non-inferiority margin recommended by the FDA for switch trials [15]. The FDA notes that most switch trials are not blinded and suggests that randomized patients discontinuing a control regimen to start the alternative regimen should not be considered as having failed on the control regimen. We followed that policy in our analyses.

In our emulated trials, there were few virological failures and few treatment changes due to toxicity. Any of these regimens would seem suitable for patients needing an NRTI-sparing regimen. The alternative regimens will be useful for patients that need to avoid dolutegravir – this might include patients with diabetes, a history of depression or women intending to conceive [33–35]. In current guidelines, raltegravir with boosted darunavir is recommended for patients that need an NRTI-sparing regimen [6]. Our two partly NRTI-sparing regimens, with lamivudine or emtricitabine, are not yet recommended but are being evaluated [6]. Our clinicians used all these regimens with caution, often adding drugs or switching back to a three-drug regimen when tenofovir alafenamide became available. Our effectiveness outcome was dominated by changes for reasons other than virological failure or toxicity. This same situation exists in real drug trials [36], so that the FDA's composite outcome is now a poor surrogate for disease progression [37] and this is one reason why long-term observational studies are needed for any new regimen.

Note that we take a time to loss of virological response approach to analysis [38] rather than the FDA's more recent snapshot approach [15]. We then estimate relative effectiveness at 48 weeks through modelling. The advantage is that outcomes beyond 48 weeks are still included in the analysis, and with few events this increases precision. The disadvantage is the modelling requires an assumption of proportional hazards. There was some attenuation in the HR as the length of follow-up increased. This implies an early effect such that the risk of regimen failure becomes increasingly similar in the two regimens the longer patients remaining on their regimen. Our estimates were similar to those reported at 48 weeks in other dolutegravir switch studies: 87% of 206 patients on dolutegravir with lamivudine [39]; 91% of 130 patients on dolutegravir with boosted darunavir [40]. However, our estimates were lower than those reported at 48 weeks for the alternative regimens: 90% of 249 patients on boosted darunavir with lamivudine [25]; 93% of 82 patients on raltegravir with boosted darunavir [27].

Poor quality observational studies have led to renewed interest in designing studies as if they were randomized trials [41]. In this context, FDA guidance to industry provides a valuable template when using observational data to emulate a trial [15,38]. Our strategy was to use this template and supplement the few data available with background clinical knowledge so that it was possible to quickly make stable estimates of relative effectiveness for clinicians needing to use novel combinations, then update estimates so that background information would be superseded by data as they accrue. Note that in this case, estimates favourable to dolutegravir and the attenuation of those estimates in the update could not be the consequence of including prior information. The priors we asserted were conservative: an HR of one is consistent with no difference between regimens and our assumed rate of events among control patients (10 per 100 patient-years) turned out to be lower than observed (33 and 27 per 100 patient-years for boosted darunavir and raltegravir controls, respectively).

Despite adopting ideas from industry, our emulated trials were not randomized. Our estimates represent subjective statistical inference, compelling only to those who are convinced of the adequacy of our conceptual trials and the appropriateness of our prior information [42]. This underscores the importance of considering alternative treatment assignments models and of updating analyses to reduce the influence of prior information. In particular, it was difficult to model the reasons why patients were started on one regimen rather than another, when we had no record of why regimens were started; only why drugs were stopped. Treatment assignment models must contain all variables necessary...
to recreate the process of assigning one regimen rather than the other; these variables must be accurately measured and represented in models using an appropriate functional form. Treatment assignment modelling suggests that our trial of fully NRTI-sparing regimens may give more reliable inference than our trial of partly NRTI-sparing regimens. However, alternative modelling of the latter led to consistent results. We did not expect our results to provide definitive answers immediately. Updating our analyses a year later illustrated the stability of early estimates although there was little gain in precision with the update because few new patients started control regimens. This suggests that it will be more difficult than expected to increase the precision of these estimates in the future.

Clinicians may now be using dolutegravir with lamivudine or emtricitabine to prevent future toxicities rather than solely to avoid existing toxicities. Under our alternative treatment allocation model, there were no matching controls for a subgroup of case patients far more likely to receive dolutegravir than boosted darunavir. Rather than use propensity scores to exclude case patients, it would be better to add trial criteria so that all trial patients had to have evidence that both abacavir and tenofovir were contraindicated. Unfortunately, when we did this, our estimates were too imprecise to be of much value.

Non-inferior effectiveness is necessary but not sufficient in a new regimen [43]. The new regimen must benefit patients: other outcomes need to be assessed to see whether these new dolutegravir regimens offer some benefit over the alternatives. The obvious outcomes are the toxicities that led to use of these regimens in the first place [8] – biomarkers of myocardial infarction, renal failure and osteoporosis – and this is another reason why long-term observational studies are still needed. In addition, the release of tenofovir alafenamide has implications for the role of NRTI-sparing dual regimens [44]. No trial has yet compared an NRTI-sparing regimen to a one pill once-a-day three-drug regimen that includes tenofovir alafenamide [8]. Soon it may be possible to emulate such a trial via a cohort collaboration. As yet, we have too few patients in the SHCS to emulate such a trial and too few patients on dolutegravir with rilpivirine to compare this regimen to an earlier equivalent (such as raltegravir with etravirine).

We used recent observational data and a structured approach to emulate clinical trials that clinicians would prefer to see but which may never happen. With our results we can quickly reassure our clinicians that the dolutegravir-based dual regimens they are using when patients need an NRTI-sparing regimen are no less effective than the dual regimens they used previously – and that those earlier regimens are also a reasonable choice if cheaper or for patients with certain conditions.

Acknowledgements

This study has been financed within the framework of the Swiss HIV Cohort Study, supported by the Swiss National Science Foundation (grant #177499), by SHCS project #832 and by the SHCS research foundation. The data are gathered by the five Swiss University Hospitals, two Cantonal Hospitals, 15 affiliated hospitals and 36 private physicians. The Basel Institute for Clinical Epidemiology and Biostatistics is supported by the Stiftung Institut für Klinische Epidemiologie.


Disclosure statement

JY has received payment from Viiv Healthcare (Viiv) for attending an advisory board meeting. AC’s institution received unrestricted educational grants from Merck Sharp and Dohme (MSD) and Viiv. PET’s institution has received research grants and advisory fees from Viiv. MC’s institution received research grants from Viiv. PV’s institution has received unrestricted research grants from Viiv and consulting or advisory board fees from MSD and Viiv. HFG has received fees for data and safety monitoring board membership from MSD. HCB has received support for travelling, consultancy fees and honorarium from MSD and Viiv. All other authors have no conflicts.

Additional file

Additional file 1: Supplementary information can be found at https://www.intmedpress.com/uploads/documents/4454_Young_Addfile1.pdf

References


