Appendix

1.   Inverse probability weights

When fitting models, the data for each patient is represented by a set of intervals, one for each month of follow up. For any given month, covariates values are as at the end of the previous month, exposure to abacavir is as at the first day of the month and the outcome (a CVD event or not) is as at the end of the month. This temporal ordering of data reflects our view of the relationships between covariates, exposure and outcome (Figure S1) and ensures that our predictors of exposure are assessed before exposure is assessed as required by inverse probability of treatment models.

Our stabilised inverse probability of treatment weights are found using eight different logistic regression models. First a denominator model is needed with which to calculate the probability an abacavir-naive patient starts treatment with abacavir given all covariates in the conventional model and all time dependent covariates potentially on the causal pathway between exposure to abacavir and cardiovascular disease (CVD). Second a numerator model is needed to calculate this probability given just baseline values of all covariates in the denominator model. These two models are fit to data up until a patient first starts abacavir (or until follow up ends for those that never start). Separate denominator and numerator models are then needed for continued use of abacavir,[1] and these models are fit to all data for patients using abacavir after use of abacavir begins. These models use the same covariates as before, with a single additional time dependent covariate – abacavir use in the previous month. This covariate takes value zero for a patient re-starting abacavir in any given month. Finally all four models are fit separately to data both before and after February 2008, because prescribing behaviour changed after the D:A:D's results were published.[2] The weight for a given patient in a given month is then calculated as the product (up to that month) of ratios of the probability of being on abacavir estimated from the appropriate numerator and denominator models respectively.[3]

In these models for abacavir use (Tables S1 and S2), patients with a high viral load, low CD4 cell count or with dyslipidaemia or lipodystrophy were more likely to start abacavir. Patients exposed to tenofovir were both less likely to start abacavir and less likely to continue its use. Patients with a previous CVD event were more likely to both start and continue using abacavir until February 2008;

after that, such patients were less likely to start abacavir and more likely to discontinue its use. Patients with a high viral load were more likely to discontinue abacavir.

Note that extreme inverse probability of treatment weights can lead to highly variable estimates.[4] Hence analyses with and without truncating inverse probability of treatment weights should be carried out to assess the impact of extreme weights on both results and scientific conclusions. In all reported results, standardised inverse probability of treatment weights have been truncated at the 1st or 99th percentile of their distribution across all person-months of follow-up if below or above this value, respectively.

We do not use censoring weights in our marginal structural models. As in the D:A:D's analyses, we assume that censoring – through either administrative censoring, patients lost to follow up or deaths unrelated to CVD – is uninformative. However, by combining inverse probability of treatment weights and inverse probability of censoring weights,[3] one could account for a competing risk of death by other causes. In an earlier analyses of didanosine, including inverse probability of censoring weights did not materially alter results.[5]

2.   Alternative weight functions

Exposure to abacavir is defined as a weighted sum of use in each past month. Exposure weights are found by estimating the components of a cubic spline that represents the relative importance of exposure at different times in the past. Hence the way exposure cumulates is free to vary with time. In most applications, exposure many years ago will not affect current risk. We assume that exposure more than four years ago would have no effect on the current risk of a CVD event.

We consider nine alternative weight functions (Figure S2). These functions differ in their degree of flexibility and in whether weights are constrained to take zero value, either at the beginning or at end of the four year interval. In Figure S2, functions in top, middle and bottom rows are estimated using a cubic spline with one, two and three internal knots respectively; hence functions in the top row are less flexible than those in the bottom row. Functions in the left column are not constrained in any way and can take values other than zero at all times; functions in the middle column are forced to take zero value at the end of four years; functions in the right column are forced to take zero value both at the beginning and at the end of the four year interval. A zero weight at the

beginning of the four year interval implies there is a lag between exposure and its effect on the current risk of an event.

We chose between these nine weight functions using a Bayesian Information Criterion adapted for censored survival data as a measure of goodness of fit.[6] In Figure S2, functions are shown for cumulative exposure modelling with both marginal structural (solid curve) and conventional (dashed curve) Cox models. The Bayesian Information Criterion is an appropriate statistic for choosing between conventional Cox models but it is not clear yet whether this is the best statistic for choosing between marginal structural models.[5,7] The weight function with the lowest value of this statistic among conventional Cox models uses a one knot spline constrained to take zero value both at the beginning and at the end of the four year interval (Figure S2, top right).

3. Abacavir and chronic kidney disease

We re-fit weighted models for the effect of exposure to abacavir with a time updated indicator of chronic kidney disease added to the covariates used to calculate inverse probability of treatment weights. We define chronic kidney disease as an estimated glomerular filtration rate below 60 ml/min/1.73m^2 (calculated using CKD-EPI equation [8]). This sensitivity analysis is of a truncated data set, limited to follow up after January 2002 when routine serum creatinine measurement began in the Swiss HIV Cohort Study (SHCS). This reduces the number of events in the data from 365 to 332. While chronic kidney disease proved to be a strong predictor of starting abacavir (data not shown), controlling for this additional confounder did not attenuate estimates of the effect of abacavir (Table S3). When re-fitting the same cumulative exposure model as before, the total effect of always being exposed to abacavir, during the entire four year period, versus never being exposed was HR 2.39 (95% CI 1.69 to 3.37).

4. Abacavir and patient subgroups

In two unplanned sensitivity analyses, we re-fit our cumulative exposure model for abacavir after excluding patients infected with HIV through injection drug use and after excluding patients exposed to abacavir before their first cardiovascular risk assessment.

The first of these sensitivity analyses is of interest because the effect of recent exposure to abacavir on the risk of myocardial infarction was attenuated in a French cohort study when drug users were

excluded [9] and there is evidence linking cocaine use to CVD in patients with HIV.[10,11] We were not able to adjust for time dependent injection drug use because routine recording of this only began in the SHCS in July 2008.

The second of these sensitivity analyses is of interest because including existing users of abacavir in an analysis may introduce bias if these patients are in some sense 'survivors' and at low risk of CVD.[12] This might result in an underestimate of a harmful exposure. The other problem created by including existing users is time dependent confounding at study entry;[12] however our use of inverse probability of treatment weights for the continued use of abacavir ought to eliminate or at least reduce such confounding.

The number of CVD events in these subgroups was reduced (from 365) to 286 in patients not infected through injection drug use and 288 in abacavir naive patients. Estimates of the effect of abacavir were not attenuated in the two analyses (Table S3). When re-fitting the same cumulative exposure model as before, the total effect of always being exposed to abacavir, during the past four years, versus never being exposed was then HR 2.54 (95% CI 1.73 to 3.73) in those not infected through injection drug use and HR 2.42 (95% CI 1.72 to 3.42) in abacavir naive patients.

Note that the analysis of abacavir naive patients provides an 'as-treated' estimate of the risk of CVD for 'new users' of abacavir.[12,13] Like a 'per-protocol' analysis, inference from this analysis is of the effect of abacavir on those who initiate and then remain on abavavir. The advantage of an 'as-treated' estimate over a 'per-protocol' estimate is that the artificial censoring required for a 'per-protocol' estimate often leads to imprecision. The 'as-treated' estimate requires that the dose-response relationship between exposure and outcome is correctly specified but our use of flexible weight functions is protection again misspecification.

The population of abacavir naive patients corresponds to the 'full population' used in an analysis recently reported by the NA-ACCORD.[14] They also analysed a 'restricted population' of antiretroviral therapy naive patients. This greatly reduced the number of events in their analysis. In these data, there are only 66 events for patients with a CVD assessment before or when first exposed to antiretroviral therapy. Experience with cumulative exposure modelling suggests we need at least 200 events before we can expect stable estimates of weight functions.

5. The effect of cumulative exposure to didanosine

Full details are available of our cumulative exposure modelling of the risk of CVD with exposure to didanosine.[5] What follows is a summary of our motivation, methods and results.

Didanosine is not a recommended component of antiretroviral therapy in US, European or World Health Organisation guidelines. However there has been a renewed interest in didanosine because in two meta-analyses, combinations with didanosine were more effective and better tolerated than other alternatives.[15,16] The wide availability and low cost, now that a generic product has been approved,[17] make this drug of interest in resources-limited settings.[18]

Clinicians prescribe didanosine for other reasons than those that motivate the use of abacavir. Hence our conventional and marginal structural Cox models for the effect of exposure to didanosine require different variables. We adjust conventional Cox models for variables identified in an earlier propensity score model for exposure to didanosine:[19] age, sex, likely transmission through injection drug use, Caucasian ethnicity, education, and time varying covariates calendar year, indicators of hepatitis infection (chronic B or C), fat loss, diabetes, nervous system toxicity, gastrointestinal toxicity, pregnancy, stages of HIV infection (CDC group A, B or C) and current use of zalcitabine and stavudine. We assume that the number of previously failed regimens and current use of tenofovir might be on a causal pathway between the use of didanosine and CVD. We do not adjust for these two variables in our conventional Cox model. For our marginal structural models, we calculate inverse probability of treatment weights using all these variables in logistic regression models, but without separate logistic regression models before and after February 2008 because we think it unlikely that the D:A:D's results changed the way clinicians prescribe didanosine.

In conventional and marginal structural Cox models, cumulative use of didanosine was not associated with an increase in the risk of a CVD event (Table S3). The hazard ratio for recent exposure in the last six months (HR 1.11, 95% CI 0.75 to 1.65) was too wide to draw conclusions about whether such exposure increases risk. Here the original D:A:D analysis had more power and the risk of a CVD event increased with recent exposure to didanosine (cumulative exposure HR 0.99, 95% CI 0.94 to 1.05; recent exposure HR 1.40, 95% CI 1.11 to 1.77 [20]). In the updated D:A:D analysis, recent but not cumulative use of didanosine was still associated with an increased risk of myocardial infarction.[21] The results of a case control study, while under-powered, suggest that recent exposure could be harmful while cumulative exposure could be protective.[9]

In weighted cumulative exposure modelling, we first assumed that exposure more than four years ago would have no effect on the current risk of a CVD event. However weight function estimates were unstable towards the end of the four year period and we reduced the assumed period of influence from 48 to 30 months. Within the first 30 months, weight estimates were similar under either assumption. This instability was probably because in our data, fewer patients were exposed to didanosine and for shorter periods relative to those exposed to abacavir. During follow up, 1876 patients were exposed to didanosine, for a median duration of 2.2 years (IQR 0.7 to 4.7) and of these, 1736 stopped taking didanosine during follow up and only 397 re-started again.

Of the nine weight functions considered, the best fitting weight function (with the lowest Bayesian Information Criterion) had a single knot and weights of zero at both the beginning and end of the 30 month interval (Figure 2, left). However this weight function suggests that exposure to didanosine has early harmful and then later protective effects. Weights assigned to exposure in the first year were positive, implying that early exposure increases the risk of a CVD event, with exposure about six months ago having the greatest effect. In contrast, didanosine use between one and two years ago was associated with negative weights, implying that such exposure reduces the risk of a CVD event. Of note, of the nine alternative weight functions considered, all but one had evidence of this dual effect assuming a 30 month period of influence, and all had evidence of this dual effect assuming a 48 month period of influence.[5]

The implications of this dual effect weight function are that the risk of a CVD event increased during the first year of uninterrupted exposure (Figure 2, right). After a year, those who continued to use didanosine had more than twice the risk of those never using didanosine (HR 2.03, 95% CI 0.74 to 5.60). If use continued, however, the risk of a CVD event declined but exposure never became protective: the total effect of always being exposed to didanosine versus never being exposed over the 30 month period was HR 0.85 (95% CI 0.47 to 1.52).

A dual effect of exposure to didanosine on the risk of a CVD event was not expected and must be viewed with caution. Didanosine is associated with peripheral neuropathy; however the risk of peripheral neuropathy does not seem to cumulate with exposure but rather peaks in the first three months of use and then subsides.[22] On the other hand, immunosuppression in advanced HIV infection is also associated with an increased risk of peripheral neuropathy.[23] Hence didanosine potentially has a dual effect on this form of neuropathy, with a short term risk of drug induced neuropathy but a long term protective effect as immunosuppression abates under continued

6

effective therapy. It is plausible that didanosine has a similar dual effect on other forms of neuropathy such as the cardiovascular autonomic neuropathy associated with sudden myocardial infarction in diabetic patients.[24] The apparent protective effect of DDI use one to two years ago may also reflect a depletion of the susceptible if most patients prone to adverse events discontinue the drug after only a short time.

6.   The effect of cumulative exposure to tenofovir

In general, clinicians prescribe tenofovir for the same reasons they prescribe abacavir – both drugs avoid or partially reverse the lipoatrophy caused by older drugs of the same class.[25] In the updated D:A:D analysis, patients exposed to tenofovir were similar to those exposed to abacavir.[21] We therefore assessed the effects of cumulative exposure to tenofovir using the same models as used for abacavir.

During follow up, 7429 patients were exposed to tenofovir, for a median duration of 3.3 years (IQR 1.5 to 5.6) and of these, 2751 stopped taking tenofovir during follow up and 1268 re-started again. There was no evidence from conventional and marginal structural Cox models that either recent or cumulative exposure to tenofovir increases the current risk of a CVD event (Table S3). In cumulative exposure modelling, as with abacavir, the best fitting weight function (with the lowest Bayesian Information Criterion) had a single knot and weights of zero at both the beginning and end of the 4 year interval (Figure 3, left). However cumulative exposure seemed if anything protective rather than harmful: the total effect of always being exposed to tenofovir versus never being exposed was HR 0.66, 95% CI 0.45 to 0.98 (Figure 3, right). Note that this protective effect was not apparent in cumulative exposure modelling with a conventional Cox model: the total effect of always being exposed to tenofovir versus never being exposed was then HR 0.97, 95% CI 0.69 to 1.37 (Figure 3, right). In other studies, neither recent nor cumulative exposure to tenofovir was associated with an increase in the risk of myocardial infarction,[21] while current exposure may be protective.[26]

7.   References

1.   Platt RW, Schisterman EF, Cole SR. Time-modified confounding. *American Journal of Epidemiology*. 2009;170:687-694.

2.   Antoniou T, Gillis J, Loutfy MR, et al. Impact of the Data Collection on Adverse Events of Anti-HIV Drugs Cohort Study on abacavir prescription among treatment-naive, HIV-infected patients in Canada. *J Int Assoc Provid AIDS Care*. 2014;13:153-159.

3.   Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000;11:561-570.

4.   Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168:656-664.

5.   Xiao Y, Abrahamowicz M, Moodie EEM, et al. Flexible marginal structural models for estimating the cumulative effect of a time-dependent treatment on the hazard: reassessing the cardiovascular risk of didanosine treatment in the Swiss HIV Cohort. *J Am Stat Assoc*. 2014;109:455-464.

6.   Volinsky CT, Raftery AE. Bayesian information criterion for censored survival models. *Biometrics*. 2000;56:256-262.

7.   Platt RW, Brookhart MA, Cole SR, et al. An information criterion for marginal structural models. *Stat Med*. 2013;32:1383-1393.

8.   Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150:604-612.

9.   Lang S, Mary-Krause M, Cotte L, et al. Impact of individual antiretroviral drugs on the risk of myocardial infarction in human immunodeficiency virus-infected patients: a case-control study nested within the French Hospital Database on HIV ANRS cohort CO4. *Arch Intern Med*. 2010;170:1228-1238.

10.   Ribaudo HJ, Benson CA, Zheng Y, et al. No risk of myocardial infarction associated with initial antiretroviral treatment containing abacavir: short and long-term results from ACTG A5001/ALLRT. *Clin Infect Dis*. 2011;52:929-940.

11.   Lai S, Fishman EK, Lai H, et al. Long-term cocaine use and antiretroviral therapy are associated with silent coronary artery disease in African Americans with HIV infection who have no cardiovascular symptoms. *Clin Infect Dis*. 2008;46:600-610.

12.   Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol*. 2003;158:915-920.

13.   Danaei G, Rodriguez LA, Cantero OF, et al. Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. *Stat Methods Med Res*. 2013;22:70-96.

14. Pallella FJ, Althoff KN, Moore RD, et al. Abacavir use and risk for myocardial infarction in the NA-ACCORD. Available at: http://www.croiconference.org/sites/default/files/posters-2015/749LB.pdf.

15. Carr A, Amin J. Efficacy and tolerability of initial antiretroviral therapy: a systematic review. *AIDS*. 2009;23:343-353.

16. Chowers M, Gottesman BS, Leibovici L, et al. Nucleoside reverse transcriptase inhibitors in combination therapy for HIV patients: systematic review and meta-analysis. *Eur J Clin Microbiol Infect Dis*. 2010;29:779-786.

17. US Food and Drug Administration. Approval of generic didanosine delayed-release capsules, 125 mg, 200 mg, 250 mg, and 400 mg. Available at: http://www.fda.gov/ForConsumers/ByAudience/ForPatientAdvocates/HIVandAIDSActivities/ucm207993.htm.

18. Laurent C, Tchatchueng Mbougua JB, Ngom Gueye NF, et al. Long-term effectiveness and safety of didanosine combined with lamivudine and efavirenz or nevirapine in antiretroviral-naive patients: a 9-year cohort study in Senegal. *Trop Med Int Health*. 2011;16:217-222.

19. Young J, Klein MB, Ledergerber B. Noncirrhotic portal hypertension and didanosine: a re-analysis. *Clin Infect Dis*. 2011;52:154-155.

20. Sabin CA, Worm SW, Weber R, et al. Use of nucleoside reverse transcriptase inhibitors and risk of myocardial infarction in HIV-infected patients enrolled in the D:A:D study: a multi-cohort collaboration. *Lancet*. 2008;371:1417-1426.

21. Worm SW, Sabin C, Weber R, et al. Risk of myocardial infarction in patients with HIV infection exposed to specific individual antiretroviral drugs from the 3 major drug classes: the data collection on adverse events of anti-HIV drugs (D:A:D) study. *J Infect Dis*. 2010;201:318-330.

22. Arenas-Pinto A, Bhaskaran K, Dunn D, et al. The risk of developing peripheral neuropathy induced by nucleoside reverse transcriptase inhibitors decreases over time: evidence from the Delta trial. *Antivir Ther*. 2008;13:289-295.

23. Ghosh S, Chandran A, Jansen JP. Epidemiology of HIV-related neuropathy: a systematic literature review. *AIDS Res Hum Retroviruses*. 2012;28:36-48.

24. Kuehl M, Stevens MJ. Cardiovascular autonomic neuropathies as complications of diabetes mellitus. *Nat Rev Endocrinol*. 2012;8:405-416.

25. Curran A, Ribera E. From old to new nucleoside reverse transcriptase inhibitors: changes in body fat composition, metabolic parameters and mitochondrial toxicity after the switch from thymidine analogs to tenofovir or abacavir. *Expert Opin Drug Saf*. 2011;10:389-406.

26. Bedimo RJ, Westfall AO, Drechsler H, et al. Abacavir use and risk of acute myocardial infarction and cerebrovascular events in the highly active antiretroviral therapy era. *Clin Infect Dis*. 2011;53:84-91.

27. Sylvestre MP, Abrahamowicz M. Flexible modeling of the cumulative effects of time-dependent exposures on the hazard. *Stat Med*. 2009;28:3437-3453.

Table S1. Denominator models used to calculate the probability an abacavir-naive patient started treatment with abacavir, either before or after February 2008, given all covariates in the conventional model and all time dependent covariates potentially on a causal pathway between exposure to abacavir and cardiovascular disease.

Covariate	Before February 2008		After February 2008	
	Odds ratio	(95% confidence interval)	Odds ratio	(95% confidence interval)
Male	0.96	(0.87, 1.07)	0.71	(0.61, 0.84)
Age (per 10 years)	1.03	(0.97, 1.09)	1.11	(1.02, 1.21)
Body mass index (kg/m^2)	0.98	(0.97, 1.00)	0.95	(0.93, 0.97)
Caucasian ethnicity	0.96	(0.84, 1.10)	0.99	(0.82, 1.20)
Current smoker	0.86	(0.77, 0.95)	0.97	(0.83, 1.14)
Transmission through injection drug use	0.91	(0.80, 1.03)	1.06	(0.86, 1.31)
Previous event [a]	1.11	(0.72, 1.72)	0.36	(0.11, 1.12)
Family history [b]	0.93	(0.81, 1.07)	0.81	(0.64, 1.03)
Diabetes mellitus [c]	1.09	(0.87, 1.37)	0.67	(0.45, 1.02)
Arterial hypertension [d]	1.26	(1.09, 1.45)	1.35	(1.10, 1.65)
Dyslipidemia [e]	1.43	(1.25, 1.64)	1.40	(1.13, 1.73)
Lipodystrophy [f]	1.33	(1.20, 1.48)	1.39	(1.18, 1.64)
Framingham risk score [g] - reference Low (<10%):				
Moderate (10-20%)	0.97	(0.84, 1.13)	0.88	(0.70, 1.11)
High (>20%)	1.21	(0.92, 1.59)	0.80	(0.50, 1.30)
Missing	1.43	(1.18, 1.72)	1.30	(0.85, 1.98)
CD4 cell count (per 100 cells / µL)	0.90	(0.88, 0.92)	0.89	(0.86, 0.92)
Log$_{10}$ HIV RNA (copies / mL)	1.07	(1.03, 1.11)	1.19	(1.12, 1.27)
Calendar year	1.07	(1.04, 1.10)	1.05	(1.03, 1.08)
Cumulative exposure to (per year):				
Didanosine	1.00	(0.96, 1.04)	1.04	(1.00, 1.09)
Stavudine	1.04	(1.00, 1.07)	0.93	(0.89, 0.99)
Zalcitabine	1.09	(0.98, 1.20)	0.92	(0.68, 1.25)
Zidovudine	1.01	(0.98, 1.05)	0.95	(0.91, 0.98)

Lamivudine	0.98	(0.95, 1.02)	1.06	(1.02, 1.10)
Tenofovir	0.86	(0.80, 0.93)	0.84	(0.80, 0.89)
Indinavir	1.04	(1.00, 1.08)	1.01	(0.95, 1.07)
Saquinavir	1.00	(0.96, 1.05)	0.92	(0.86, 1.00)
Nelfinavir	1.00	(0.97, 1.04)	1.01	(0.96, 1.06)
Ritonavir	1.05	(1.00, 1.09)	1.05	(1.00, 1.11)
Lopinavir	0.98	(0.92, 1.04)	1.03	(0.99, 1.08)
Amprenavir	0.66	(0.52, 0.84)	0.80	(0.64, 1.01)
Atazanavir	0.96	(0.86, 1.07)	1.01	(0.94, 1.07)
Efavirenz	1.02	(0.98, 1.06)	0.95	(0.91, 0.99)
Nevirapine	0.97	(0.92, 1.03)	0.92	(0.87, 0.98)

[a] Cardiovascular disease event before the patients first cardiovascular risk assessment.

[b] Myocardial infarction or stroke before the age of 50 in any first degree relative.

[c] Clinical diagnosis, or casual plasma glucose >11.1 mmol/L, or on anti-diabetic medication or insulin.

[d] Systolic blood pressure > 140 mmHg, diastolic blood pressure > 90 mmHg, or on anti-hypertensive medication.

[e] Total cholesterol > 6.2 mmol/L, HDL cholesterol < 0.9 mmol/L or on lipid-lowering medication.

[f] Patient and clinician report either body fat loss or body fat gain.

[g] Estimated risk of cardiovascular disease in the next 10 years.

Table S2. Denominator models used to calculate the probability a patient already exposed to abacavir continued treatment with abacavir, either before or after February 2008, given all covariates in the conventional model and all time dependent covariates potentially on a causal pathway between exposure to abacavir and cardiovascular disease.

Covariate	Before February 2008		After February 2008	
	Odds ratio	(95% confidence interval)	Odds ratio	(95% confidence interval)
Previous use of abacavir	4322	(3938, 4742)	24031	(20551, 28100)
Male	1.32	(1.20, 1.45)	1.03	(0.91, 1.18)
Age (per 10 years)	1.06	(1.00, 1.11)	1.04	(0.97, 1.12)
Body mass index (kg/m^2)	1.00	(0.99, 1.01)	1.00	(0.98, 1.01)
Caucasian ethnicity	0.91	(0.80, 1.03)	1.18	(1.01, 1.38)
Current smoker	0.91	(0.83, 0.99)	0.88	(0.78, 1.00)
Transmission through injection drug use	0.82	(0.74, 0.91)	0.98	(0.84, 1.15)
Previous event [a]	1.06	(0.72, 1.56)	0.39	(0.24, 0.63)
Family history [b]	0.87	(0.77, 0.98)	0.89	(0.75, 1.06)
Diabetes mellitus [c]	0.99	(0.82, 1.18)	0.84	(0.68, 1.04)
Arterial hypertension [d]	1.08	(0.96, 1.23)	0.84	(0.73, 0.96)
Dyslipidemia [e]	1.76	(1.56, 1.99)	1.10	(0.95, 1.28)
Lipodystrophy [f]	0.88	(0.81, 0.96)	0.90	(0.79, 1.01)
Framingham risk score [g] - reference Low (<10%):				
Moderate (10-20%)	0.86	(0.77, 0.97)	0.86	(0.74, 1.01)
High (>20%)	0.80	(0.63, 1.02)	0.74	(0.56, 0.98)
Missing	14	(12, 17)	100	(67, 149)
CD4 cell count (per 100 cells / µL)	1.04	(1.02, 1.05)	1.04	(1.02, 1.06)
Log$_{10}$ HIV RNA (copies / mL)	0.62	(0.60, 0.65)	0.66	(0.61, 0.71)
Calendar year	1.14	(1.11, 1.19)	1.15	(1.13, 1.18)
Cumulative exposure to (per year):				
Didanosine	0.99	(0.97, 1.02)	1.03	(1.00, 1.05)
Stavudine	1.00	(0.97, 1.02)	1.00	(0.97, 1.03)
Zalcitabine	1.08	(1.00, 1.16)	1.06	(0.94, 1.19)

Zidovudine	1.00	(0.98, 1.02)	0.99	(0.97, 1.01)
Lamivudine	0.99	(0.97, 1.01)	1.05	(1.03, 1.07)
Tenofovir	0.84	(0.80, 0.87)	0.91	(0.89, 0.94)
Indinavir	1.06	(1.02, 1.09)	1.02	(0.98, 1.06)
Saquinavir	1.04	(1.00, 1.08)	1.02	(0.98, 1.07)
Nelfinavir	0.99	(0.97, 1.02)	1.03	(1.00, 1.07)
Ritonavir	1.00	(0.97, 1.03)	1.03	(0.99, 1.07)
Lopinavir	0.94	(0.91, 0.97)	1.00	(0.97, 1.02)
Amprenavir	1.06	(1.01, 1.11)	0.98	(0.93, 1.03)
Atazanavir	0.86	(0.80, 0.92)	0.98	(0.94, 1.03)
Efavirenz	1.01	(0.98, 1.03)	1.03	(1.01, 1.05)
Nevirapine	0.96	(0.93, 0.99)	1.03	(1.00, 1.06)

[a] Cardiovascular disease event before the patients first cardiovascular risk assessment.

[b] Myocardial infarction or stroke before the age of 50 in any first degree relative.

[c] Clinical diagnosis, or casual plasma glucose >11.1 mmol/L, or on anti-diabetic medication or insulin.

[d] Systolic blood pressure > 140 mmHg, diastolic blood pressure > 90 mmHg, or on anti-hypertensive
medication.

[e] Total cholesterol > 6.2 mmol/L, HDL cholesterol < 0.9 mmol/L or on lipid-lowering medication.

[f] Patient and clinician report either body fat loss or body fat gain.

[g] Estimated risk of cardiovascular disease in the next 10 years.

Table S3. The relative risk of a cardiovascular disease event in additional analyses for patients exposed to abacavir, didanosine or tenofovir.

Exposure parameters	Conventional model		Marginal structural model	
	Hazard ratio	(95% CI)	Hazard ratio	(95% CI)
Abacavir (allowing for chronic kidney disease) [a]				
Cumulative exposure (per year)	1.04	(0.99, 1.10)	1.04	(0.98, 1.10)
Recent exposure within past 0 to 6 months	1.54	(1.13, 2.10)	1.61	(1.15, 2.25)
Abacavir (in patients not infected by drug use) [b]				
Cumulative exposure (per year)	1.06	(1.00, 1.13)	1.04	(0.97, 1.12)
Recent exposure within past 0 to 6 months	1.44	(1.04, 2.01)	1.73	(1.16, 2.60)
Abacavir (in abacavir naive patients) [c]				
Cumulative exposure (per year)	1.09	(1.01, 1.17)	1.08	(1.00, 1.17)
Recent exposure within past 0 to 6 months	1.46	(1.02, 2.09)	1.44	(0.98, 2.12)
Didanosine [d]				
Cumulative exposure (per year)	1.00	(0.94, 1.06)	1.04	(0.95, 1.13)
Recent exposure within past 0 to 6 months	1.11	(0.75, 1.65)	0.89	(0.50, 1.57)
Tenofovir [e]				
Cumulative exposure (per year)	1.01	(0.94, 1.10)	0.94	(0.86, 1.03)
Recent exposure within past 0 to 6 months	0.94	(0.69, 1.29)	0.86	(0.60, 1.22)

[a] Data limited to follow up after January 2002 when routine serum creatinine measurement began. Conventional model adjusted as in Table 2. Marginal structural model fit as in Table 2 except with a time updated indicator of chronic kidney disease added to the covariates used to calculate inverse probability of treatment weights.

[b] Equivalent models to those in Table 2 but excluding patients most likely infected with HIV through injection drug use.

[c] Equivalent models to those in Table 2 but excluding patients exposed to abacavir before their first cardiovascular risk assessment.

[d] Conventional model adjusted for age, sex, likely transmission through injection drug use, Caucasian ethnicity, education, and time varying covariates calendar year, indicators of hepatitis infection (chronic B or C), fat loss, diabetes, nervous system toxicity, gastrointestinal toxicity, pregnancy, stages of HIV infection (CDC group A, B or C) and current use of zalcitabine and stavudine. Marginal structural model fit using inverse probability of treatment weights, with weights found in four different logistic regression models. The covariates in these models are those used in the conventional models plus time varying indicators for the number of previously failed regimens and current use of tenofovir. Both models are fit to the full data set, as in Table 2.

[e] Equivalent models to those in Table 2 and fit to the full data set, as in Table 2.

Figure S1. A directed acyclic graph (DAG) showing assumed temporal relationships between time-varying covariates (V_t), exposure to abacavir (A_t) and a possible cardiovascular disease (CVD) event (Y_t) in the three months after a first cardiovascular risk assessment (t=0, 1, 2, 3). Some time-varying risk factors for CVD (such as smoking status and body mass index) are thought to not lie on the causal pathway between exposure to abacavir and CVD – for these covariates, there will be no arrows between A_t and V_{t+1}.
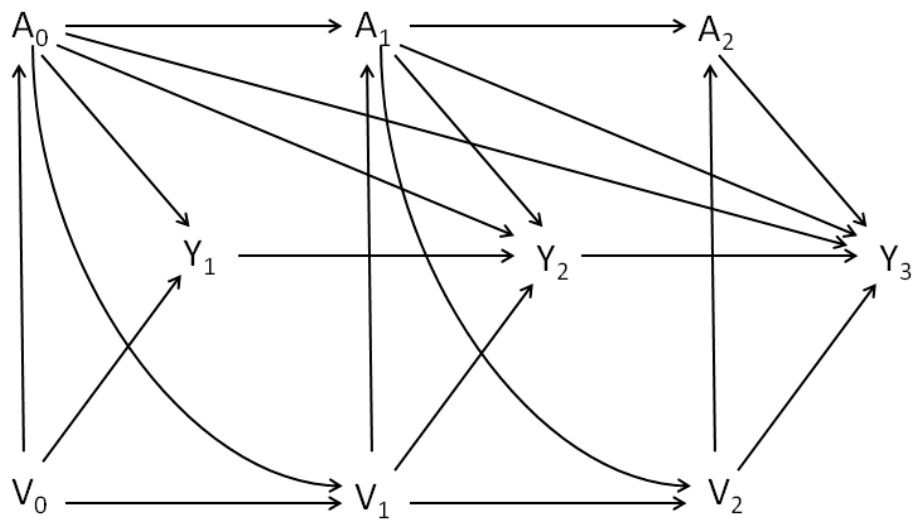
Figure S2. All 9 estimated weight functions for the effect of exposure to abacavir on the risk of cardiovascular disease events. Exposure more than four years ago was assumed to have no effect on current risk. Functions in top, middle and bottom rows are estimated using a cubic spline with one, two and three internal knots respectively. Functions in the left column are not constrained in any way and can take values other than zero at all times; functions in the middle column are forced to take zero value at the end of four years; functions in the right column are forced to take zero value both at the beginning and at the end of the four year interval. Functions are shown for cumulative exposure modelling with both marginal structural (solid curve) and conventional (dashed curve) Cox models [5,27] and the Bayesian Information Criterion (BIC) is given for each function as a measure of goodness of fit.[6]