Appendix

In this appendix, we describe the inverse probability weights we used for confounder control and give the results of three sensitivity analyses.

Inverse probability weights

The following two sets of weights were constructed to adjust for differences in the sort of patients starting each of two therapies and in the sort of patients remaining on each therapy over time [1].

First we constructed point treatment weights to adjust for differences in the sort of patients starting each of two therapies. We fit a logistic regression model to describe the probability that a patient received one therapy rather than the other, and included as predictor variables: gender, ethnicity, intravenous drug use as the likely mode of HIV-transmission, duration of HIV infection, age, advanced HIV-infection (CDC group C), diabetes, hypertension, infection with chronic hepatitis B or hepatitis C, viral load, and $CD4^+$ T-cell count. The point treatment weight for each patient was the inverse of the probability of receiving the therapy the patient actually started, stabilised using the proportion of patients receiving that therapy [2].

Second we constructed time-dependent censoring weights to adjust for differences between the patients who remained in the analysis and the patients who were censored either because of administrative censoring, because they dropped out of the SHCS or because they stopped taking any one of the four drugs tenofovir, efavirenz, LPV/r or ATV/r. We treated each person-month as an observation, and fit a pooled logistic regression model for the probability of remaining uncensored through to the month censoring occurred. In this model, we included a time-dependent spline intercept for the months since starting therapy [3] and as predictor variables: efavirenz or PI based therapy, gender, ethnicity, intravenous drug use as the likely mode of HIV-transmission, age, duration of HIV infection when starting therapy, and time-updated variables for advanced HIV-infection (CDC group C), diabetes, hypertension, infection with chronic hepatitis B or hepatitis C, eGFR, $CD4^+$ T-cell count, and virological failure (either incomplete suppression or virological rebound after suppression – see [4]). The correct temporal order between predictor variables and the response (remaining uncensored) was maintained by using time-updated values of predictors from the previous follow up visit. A missing value for a

time-updated predictor was imputed by carrying forward the most recently recorded value. If a first eGFR was missing for a patient, we imputed this first value using a multiple regression model with all the other variables listed above as predictors. The censoring weight for each person-month was the inverse of the probability of remaining uncensored through to that month, stabilised using the proportion of patients on that therapy still not censored that month [5].

In an additional sensitivity analysis suggested by a reviewer, we replaced intravenous drug use as the likely mode of HIV-transmission with current intravenous drug use when constructing both point of treatment and time dependent censoring weights. If current intravenous drug use was not known for a patient prior to starting therapy, we assumed a value based on whether or not intravenous drug use was the likely mode of HIV-transmission.

In weighted analyses, we fit marginal structural models for repeated measures using generalized estimating equations with an independent working covariance matrix and with each measurement weighted by the product of its point of treatment and censoring weights. The parameters of this model will estimate the average causal effect of one therapy relative to another under the following assumptions: (1) the model used to create point of treatment weights is correctly specified with no unmeasured confounding; (2) the model used to create censoring weights is correctly specified so that censoring is ignorable once we adjust for any selection biases that arise through loss of follow up; and (3) the model describing the difference in eGFR between therapies is correctly specified.

Sensitivity analyses

Estimated differences between therapies appeared robust to outliers. After excluding values of eGFR below 30 mL/min/1.73m^2, the estimated difference in eGFR during the first 6 months of therapy was -2.9 (95% CI -7.7 to 1.9) mL/min/1.73m^2 for patients on tenofovir and LPV/r (3 values excluded), and -7.6 (95% CI -11.9 to -3.4) mL/min/1.73m^2 for patients on tenofovir and ATV/r (2 values excluded), both relative to patients on tenofovir and efavirenz.

Four patients recorded values of eGFR below 30 mL/min/1.73m^2. Two patients recorded extremely low values (<5 mL/min/1.73m^2) immediately prior to starting therapy. Both started therapy with a high viral load (above 5 $\log_{10}$ copies/mL) and a low CD4$^+$ T-cell count (below 200 cells/mm^3); one patient was hospitalised for two days. Both patients – one starting tenofovir and efavirenz, the other starting tenofovir and LPV/r – then recorded a series of eGFR

all in excess of 60 mL/min/1.73m^2 until administrative censoring after more than 3 years on therapy. A third patient started tenofovir and LPV/r, two weeks later recorded a first eGFR of 21 mL/min/1.73m^2 and immediately stopped taking tenofovir. This patient has only been on therapy for 6 months with recent values of eGFR in the range from 46 to 62. A fourth patient started tenofovir and ATV/r and after more than 3 years on therapy recorded consecutive values of 26 and 24 mL/min/1.73m^2. This patient then stopped taking tenofovir and recorded a series of improving values over the next year with the last three values of eGFR all 60 mL/min/1.73m^2 or more.

Compared to the main analysis, results based on both calibrated and uncalibrated serum creatinine measurements had slightly larger differences in eGFR between therapies during the first 6 months and slightly narrower confidence intervals (Appendix Table A1). Using both calibrated and uncalibrated serum creatinine measurements, the estimated difference in eGFR during the first 6 months of therapy was -4.4 (95% CI -8.6 to –0.2) mL/min/1.73m^2 for patients on tenofovir and LPV/r, and -8.8 (95% CI -12.4 to -5.2) mL/min/1.73m^2 for patients on tenofovir and ATV/r, both relative to patients on tenofovir and efavirenz.

Other measures of the glomerular filtration rate led to a similar pattern of differences between therapies, with differences apparent only during the first 6 months of therapy and during this period, greater differences with tenofovir and ATV/r than with tenofovir and LPV/r (Appendix Table A2). Using the Modification of Diet in Renal Disease (MDRD) Study equation [6], the estimated difference in eGFR during the first 6 months of therapy was -2.6 (95% CI -8.5 to 3.2) mL/min/1.73m^2 for patients on tenofovir and LPV/r, and -10.0 (95% CI -14.3 to -5.6) mL/min/1.73m^2 for patients on tenofovir and ATV/r, both relative to patients on tenofovir and efavirenz.

In the additional sensitivity analysis suggested by a reviewer, the estimated difference in eGFR for LPV/r based therapy was -3.1 (95% CI -7.8 to 1.6) mL/min/1.73m^2 during the first 6 months of therapy, then followed by a difference of 0.2 (95% CI -0.8 to 1.2) mL/min/1.73m^2 for each additional 6 months of therapy; for ATV/r based therapy, the difference was -7.6 (95% CI -11.8 to -3.4) mL/min/1.73m^2 during the first 6 months of therapy, then followed by a difference of -0.5 (95% CI -1.6 to 0.6) mL/min/1.73m^2 for each additional 6 months of therapy.

References

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

2. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; **11**:550-560.

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. Young J, Bucher HC, Guenthard HF, Rickenbach M, Fux CA, Hirschel B, *et al*. Virological and immunological responses to efavirenz or boosted lopinavir as first-line therapy for patients with HIV. *Antivir Ther* 2009; **14**:771-779.

5. Hernan MA, Brumback BA, Robins JM. Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. *Stat Med* 2002; **21**:1689-1709.

6. Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, *et al*. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. *Ann Intern Med* 2006; **145**:247-254.

7. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, III, Feldman HI, *et al*. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009; **150**:604-612.

8. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. *Nephron* 1976; **16**:31-41.

Table A1: Differences in average estimated glomerular filtration rate (eGFR)[1] found from calibrated serum creatinine measurements only (the main analysis) or from both calibrated and uncalibrated serum creatinine measurements. Patients starting therapy with tenofovir and ritonavir boosted lopinavir (LPV/r) or tenofovir and ritonavir boosted atazanavir (ATV/r) are both compared to patients starting tenofovir and efavirenz.

Difference in eGFR[1] per 6 months (95% CI), relative to tenofovir and efavirenz (n=484 or 712)		Tenofovir and LPV/r (n=269 or 308)	Tenofovir and ATV/r (n=187 or 287)
Unadjusted model			
≤ 6 months	Calibrated measurements only	-4.6 (-8.6 to -0.5)	-7.2 (-11.3 to -3.2)
	All measurements	-6.3 (-9.9 to -2.7)	-8.8 (-12.2 to –5.4)
> 6 months	Calibrated measurements only	0.3 (-0.4 to 1.1)	-0.4 (-1.6 to 0.8)
	All measurements	0.7 (0.0 to 1.4)	-0.4 (-1.4 to 0.6)
Weighted model			
≤ 6 months	Calibrated measurements only	-2.6 (-7.3 to 2.2)	-7.6 (-11.8 to -3.4)
	All measurements	-4.4 (-8.6 to -0.2)	-8.8 (-12.4 to –5.2)
> 6 months	Calibrated measurements only	-0.0 (-1.1 to 1.1)	-0.5 (-1.6 to 0.7)
	All measurements	-0.4 (-0.4 to 1.2)	-0.2 (-1.2 to 0.9)

Abbreviations: LPV/r, ritonavir-boosted lopinavir; ATV/r, ritonavir-boosted atazanavir; CI, confidence interval.

[1] Estimated glomerular filtration rate (mL/min/1.73m^2,) calculated with the Chronic Kidney Disease Collaboration (CKD-EPI) equation [7].

Table A2: Differences in the glomerular filtration rate calculated with the Chronic Kidney Disease Collaboration (CKD-EPI) equation (the main analysis) or with other alternative equations. Patients starting therapy with tenofovir and ritonavir boosted lopinavir (LPV/r) or tenofovir and ritonavir boosted atazanavir (ATV/r) are both compared to patients starting tenofovir and efavirenz.

Difference in renal function per 6 months (95% CI), relative to tenofovir and efavirenz (n=484)		Tenofovir and LPV/r (n=269)	Tenofovir and ATV/r (n=187)
Unadjusted model			
≤ 6 months	CKD-EPI equation[1]	-4.6 (-8.6 to -0.5)	-7.2 (-11.3 to -3.2)
	MDRD Study equation[2]	-5.2 (-9.8 to -0.5)	-10.2 (-14.7 to -5.7)
	Cockcroft-Gault equation[3]	-5.8 (-12.6 to 1.1)	-6.9 (-13.8 to 0.0)
> 6 months	CKD-EPI equation[1]	0.3 (-0.4 to 1.1)	-0.4 (-1.6 to 0.8)
	MDRD Study equation[2]	-0.1 (-1.0 to 0.8)	0.0 (-1.6 to 1.6)
	Cockcroft-Gault equation[3]	0.6 (-0.9 to 2.2)	0.2 (-1.8 to 2.1)
Weighted model			
≤ 6 months	CKD-EPI equation[1]	-2.6 (-7.3 to 2.2)	-7.6 (-11.8 to -3.4)
	MDRD Study equation[2]	-2.6 (-8.5 to 3.2)	-10.0 (-14.3 to -5.6)
	Cockcroft-Gault equation[3]	-2.1 (-9.1 to 4.9)	-8.4 (-14.6 to -2.2)
> 6 months	CKD-EPI equation[1]	-0.0 (-1.1 to 1.1)	-0.5 (-1.6 to 0.7)
	MDRD Study equation[2]	-0.4 (-1.5 to 0.7)	-0.3 (-1.6 to 0.9)
	Cockcroft-Gault equation[3]	0.4 (-1.5 to 2.2)	0.0 (-1.5 to 1.5)

Abbreviations: LPV/r, ritonavir-boosted lopinavir; ATV/r, ritonavir-boosted atazanavir; CI, confidence interval.

[1] Estimated glomerular filtration rate (mL/min/1.73m^2) calculated with the Chronic Kidney Disease Collaboration (CKD-EPI) equation [7].

[2] Estimated glomerular filtration rate (mL/min/1.73m^2) calculated with the Modification of Diet in Renal Disease (MDRD) Study equation [6].

[3] Creatinine clearance (mL/min) calculated with Cockcroft-Gault (CG) equation [8]. Missing weight measurements reduce the number of patients in these analysis to n=482, n=267 and n=185 for patients starting tenofovir with efavirenz, LPV/r and ATV/r respectively.