Tropical Medicine and International Health doi:10.1111/j.1365-3156.2010.02507.x

VOLUME 15 SUPPL. 1 PP 63–69 JUNE 2010

Tracking a sample of patients lost to follow-up has a major impact on understanding determinants of survival in HIV-infected patients on antiretroviral therapy in Africa

Elvin H. Geng[1,2], David V. Glidden[3,2], Nneka Emenyonu[4,2], Nicolas Musinguzi[4,2], Mwebwesa Bosco Bwana[4,2], Torsten B. Neilands[5,2], Winnie Muyindike[4,2], Constantin T. Yiannoutsos[6,2], Steven G. Deeks[1,2], David R. Bangsberg[4,7,2] and Jeffrey N. Martin[1,3,2]

1 *Division of HIV/AIDS, Department of Medicine, San Francisco General Hospital, USA*
2 *East Africa International Epidemiologic Databases to Evaluate AIDS (IeDEA) Consortium*
3 *Department of Epidemiology and Biostatistics, University of California, San Francisco, USA*
4 *Mbarara University of Science and Technology, Mbarara, Uganda*
5 *Center for AIDS Prevention Studies, University of California, San Francisco, USA*
6 *Division of Biostatistics, Department of Medicine, Indiana University, Indianapolis IN, USA*
7 *Massachusetts General Hospital Center for Global Health, Harvard Medical School, Boston MA, USA*

Summary

OBJECTIVE To date, data regarding the determinants of mortality in HIV-infected patients starting antiretroviral therapy (ART) in Africa have been primarily derived from routine clinical care settings practicing the public health approach. Losses to follow-up, however, are high in these settings and may lead to bias in understanding the determinants of mortality.

METHODS We evaluated HIV-infected adults initiating ART between January 1, 2004 and September 30th, 2007 in an ART clinic in southwestern Uganda. Clinical and demographic characteristics were obtained through routine clinical care. In evaluating determinants of mortality, a 'naïve' analysis used only deaths known through routine processes. A 'sample-corrected' approach incorporated, through probability weights, outcomes from a representative sample of patients lost to follow-up whose vital status was ascertained through tracking in the community.

RESULTS In 3,628 patients followed for up to 3.75 years after ART initiation, the 'naïve' approach identified male sex and lower pre-ART CD4 count as independent determinants of mortality. The 'sample-corrected' approach found lower pre-ART CD4 count, older age, lower weight and calendar year of ART initiation, but not male sex, to be independent determinants of mortality.

CONCLUSIONS Analyses to identify determinants of mortality in HIV-infected patients on ART in Africa that do not account for losses to follow-up can identify spurious associations and miss actual relationships – both with the potential to mislead public health efforts. A sampling-based approach to account for losses to follow-up represents a feasible and potentially scalable method to strengthen the evidence available for implementation of ART delivery in Africa.

keywords antiretroviral scale-up strategies, losses to follow-up, determinants of mortality on antiretroviral therapy, sampling studies, monitoring and evaluation, Africa

Introduction

As global access to antiretroviral therapy (ART) for HIV-infected patients expands rapidly, optimizing outcomes of patients starting complicated, potentially toxic and lifelong regimens in resource limited settings is an emerging public health priority. Because of the scale and speed of the global roll-out, observational studies originating from routine clinical care in settings practicing the 'public health approach' (Gilks *et al.* 2006) will necessarily provide the first and largest source of evidence to guide treatment strategies. The quality of this evidence is therefore crucial for optimizing outcomes among millions of HIV-infected patients starting ART.

Among the various threats to the validity of the evidence originating from sites practicing routine delivery of ART in Africa, losses to follow-up are increasingly perceived as being among the most important (Rosen *et al.* 2007). Losses to follow-up occur in a high proportion of patients initiating ART (Braitstein *et al.* 2006; Wools-Kaloustian

et al. 2006; Rosen *et al.* 2007), and it is now well appreciated that these losses lead to substantial underestimation of mortality (Braitstein *et al.* 2006; Geng *et al.* 2008; Yiannoutsos *et al.* 2008). It is likely, furthermore, that these losses are differential on both determinants of interest (e.g. pre-therapy CD4 + T-cell count) and outcomes (e.g. mortality) and therefore represent a form of selection bias (Rothman & Greenland 2008). Therefore, equally important, but perhaps less well appreciated, is that losses to follow-up can lead to invalid findings in analyses to identify determinants of mortality. The direction of the bias is unpredictable: both spurious determinants can be identified and actual relationships can be missed – both with the potential to mislead public health efforts. Losses to follow-up, therefore, represent a serious threat to the integrity of scientific evidence guiding the global delivery of ART.

Previously, we presented a sampling-based approach to address losses to follow-up among HIV-infected patients on ART in Uganda (Geng *et al.* 2008) and Kenya (Yiannoutsos *et al.* 2008). Although finding outcomes in all patients who are lost to follow-up is frequently prohibitive because lost patients are too numerous, we demonstrated that determining outcomes in a representative *sample* of lost patients is feasible. Outcomes in this sample of patients lost to follow-up were used as an unbiased estimate of outcomes in all lost patients. These outcomes were incorporated into analyses through simple probability weights that can be implemented in standard statistical software. In our initial application of this approach, we found that sampling-based correction increased the cumulative incidence of mortality following ART initiation nearly five fold from 1.7% to 7.5% at 1 year and from 2.3% to 12.2% at 3 years (Geng *et al.* 2008). We now demonstrate the effect of this sampling-based approach on analyses to identify determinants of mortality among HIV-infected patients initiating ART in a prototypical ART delivery program in Uganda.

Methods

Patients

We have previously described the patient population and tracking procedure in more detail (Geng *et al.* 2008, 2009). In short, we evaluated all HIV-infected adults starting their first ART regimen at the Immune Suppression Syndrome (ISS) Clinic in the rural district of Mbarara, Uganda, between January 1, 2004 and September 30, 2007. Each month, patients who had been absent for 6 months or more were identified as being lost to follow-up and subsequently a consecutive and unselected sample of these patients was sought in the community by the patient tracker. Specifically, we used the clinic's electronic medical record system on the OpenMRS platform – which through internal audits captures 99.9% of all patients visits – to generate a list of patients who were lost to follow-up each month along with information on their identity (name, sex, age and occupation) and residence (district, county, sub-county, parish and village). A tracker, who was a health educator at the clinic, was chosen because of his familiarity with the surrounding community and extensive experience with HIV counseling and education and issues of confidentiality. The tracker sought to locate and directly speak to the missing patient. Failing that, he would record vital status information from close informants (e.g. family members, neighbours, or friends). As the district is mostly rural, tracking was carried out on motorcycle or on foot. The tracker sought each missing patient for up to 2 days. He was instructed to attempt to locate as many patients on the list as time permitted each month. The representativeness of the sample depends on the fraction of patients in the sample who have their vital status updated, and therefore the protocol emphasized maximizing this fraction. Given that time was limited relative to the large number of lost patients, each month only a sample of patients was sought.

Information available to the tracker comprised identifier (name, sex, age and occupation) and residence (district, county, sub-county, parish, village). The tracker sought to locate and directly speak to the missing patient, but failing that, would obtain updated vital status from close informants (e.g., family members, neighbours or friends). As the district is mostly rural, tracking was carried out on motorcycle or on foot.

Measurements

Clinical and demographic characteristics of interest were obtained from the clinic's electronic medical record system and included age, sex, weight, pre-ART CD4+T-cell count, distance from clinic to home and year of ART initiation. Age was handled as a continuous variable and rescaled to represent each 10 year increment. We chose a cut-point of 40 kg for weight because given the average height in the clinic of 160 cm (data not shown) this weight is equivalent to a body mass index (BMI) of 16 which meets the established definition of severe malnutrition (Ferro-Luzzi *et al.* 1992). The pre-ART CD4 + T-cell measurement was defined as the last determination within 6 months of ART initiation. Categories for pre-ART CD4 + T-cell values were selected according to convention at ≤50, 51–100, 101–200 and >200 cells/mm^3. Missing data were addressed with multiple imputation using the ICE and MIM (version 1.4.6) procedures in Stata version 10.0 (College

Station, TX, USA). The pre-ART CD4 + T-cell count was the variable with the most missing data, and we assumed that the probability of a CD4 + T-cell determination being missing was not dependent on its value after controlling for WHO clinical stage and calendar time. This is known as the 'missing at random' (MAR) assumption (Schafer 1999), which is justified in this context by our observation that (i) missing pre-therapy CD4 + T-cell values were commonly the result of technical limitations, which occurred at discrete calendar times, and (ii) CD4 + T-cell determinations were at times not obtained in patients with clearly advanced WHO clinical stage. The imputation model for missing CD4 + T-cell values therefore included WHO clinical stage and calendar time as well as other available covariates.

The outcome of deaths was ascertained in two ways. First, we recorded deaths known by the clinic during the course of routine care through query of the clinic's electronic medical record system. Second, deaths among the sample of sought after patients who were lost to follow-up were obtained via interview of close informants by the patient tracker.

Analyses

We evaluated determinants of mortality following ART initiation in two analyses. In the first approach, the 'naïve' analysis, we used only the deaths recorded during the course of routine clinical care. In the second approach, the 'corrected' analysis, we used both deaths recorded during the course of routine clinical care and those uncovered through the tracking of a consecutive and unselected sample of lost patients in the community. In both approaches, proportional hazards regression was used, where time zero was the date of ART initiation. Persons not known to have died were censored at the last date they were known to be alive either at clinic or in the community. In the corrected analysis, the lost patients without updated vital status ascertainment were dropped and those with updated vital status ascertainment through tracking were used to represent them with the use of a probability weight. This approach is based on the premise that the distribution of a characteristic in the sample is an unbiased estimate of the distribution of characteristic in the underlying population of patients lost to follow-up. The probability weight is the ratio of all patients lost to follow-up to those lost with subsequent vital status ascertainment through tracking. This method is equivalent to the double-sampling method of Frangakis and Rubin (2001) and yields unbiased point estimates and standard errors in the proportional hazards model (Lin 2000). Each of the seven clinical and demographic characteristics described earlier were evaluated for their association with time to death in unadjusted and multivariable analyses. In the multivariable analysis, we chose to include all seven predictor variables in the model and did not attempt to identify the most parsimonious model. This was performed to allow for a straightforward comparison of the sample-corrected and naïve analyses as well as the *a priori* importance of each of the variables. The assumption of proportional hazards was checked with graphical analysis of scaled Schoenfeld residuals. All analyses were conducted in Stata version 10.0 (College Station).

The study was approved by the institutional review board of University of California, San Francisco and the Mbarara University of Science and Technology.

Results

Characteristics of the 3628 HIV-infected adults newly initiating ART between January 1, 2004 and September 30, 2007 have been previously reported. In short, the median age was 35 years (interquartile range (IQR) 30 to 42), and 61% were women. The median CD4 + T-cell count prior to therapy in 1674 patients in whom it was available was 95/mm^3 (IQR 36 to 172). The median time between last pre-therapy CD4 + T-cell determination and ART initiation was 41 days (IQR 15 to 83).

As previously reported, over a maximum of 3.75 years of follow-up, 829 patients became lost to follow-up as defined by 6 months of absence from clinic for a cumulative incidence at 1, 2 and 3 years of 16% (95% confidence interval, 15 to 17%), 30% (95% CI, 28 to 32%), and 39% (37 to 42%), respectively. Of the 829 patients lost to follow-up, a sample of 128 (15%) was sought after in the community. Of the 128 lost patients in this sample, 111 (87%) had ascertainment of vital status. Of these 111, 32 had died. As previously reported, in 48 cases (43%), the patient was directly found, and in 63 cases (57%) an informant was found. Informants included parents (15%), children (9%), friends (5%), neighbors (8%), siblings (5%), spouse (4%), and others (11%).

The cumulative incidence of death in the lost patients at 1, 3, 6 and 12 months was 9.1% (95% CI, 5.0–16%), 15% (95% CI, 9.3 to 23%), 23% (95% CI, 16 to 33%) and 36% (95% CI, 24 to 44%), respectively. Patients who were lost and subsequently had their vital status ascertained by tracking (n = 111) were similar in age (median 35 *vs.* 36 years), gender (59% *vs.* 58% women), pre-therapy CD4 + T-cell count (median 74.5 *vs.* 72.0 cells/mm^3), to those who were lost and did not have ultimate vital status ascertained (n = 718).

Naïve analyses (Table 1), where only deaths known passively to the clinic were considered, found statistically

Table 1 Naïve analysis of factors associated with mortality among ART initiators in Mbarara, Uganda. Results of unadjusted and multivariable proportional hazards regression models using only deaths passively ascertained in routine clinical care

	Unadjusted			Multivariable		
Factor	HR	*P*-value	95% CI	HR	*P*-value	95% CI
Age (per 10 years)	1.22	0.15	0.93–1.61	1.18	0.29	0.87–1.59
Male sex	2.19	<0.01	1.30–3.72	1.86	0.03	1.06–3.26
Weight < 40 kg	1.63	0.25	0.70–3.81	1.64	0.28	0.67–4.03
Pre-ART CD4 count						
≤50 cells/mm^3	Ref			Ref		
51–100 cells/mm^3	0.49	0.08	0.22–1.09	0.56	0.15	0.25–1.24
101–200 cells/mm^3	0.19	<0.01	0.04–0.52	0.24	<0.01	0.08–0.68
>200 cells/mm^3	0.12	0.01	0.03–0.63	0.15	<0.01	0.03–0.67
Distance from home to clinic (per 10 km)	0.95	0.29	0.87–1.03	0.93	0.10	0.85–1.01
WHO clinical stage 4	1.75	0.24	0.67–4.57	1.25	0.65	0.47–3.28
Year of ART initiation						
2004	Ref			Ref		
2005	0.54	0.05	0.30–0.99	0.69	0.25	0.37–1.29
2006	0.33	0.06	0.15–0.72	0.50	0.09	0.22–1.10
2007	0.14	0.01	0.03–0.64	0.27	0.09	0.06–1.22

ART, antiretroviral therapy; HR, hazard ratio.

significant associations between death and male sex and pre-ART CD4 + T-cell count. In multivariable analyses, men had a 1.86-fold higher rate of death (95% CI, 1.06–3.26) compared to women and those with higher pre-ART CD4 + T-cell counts had lower rates of death. Compared to a reference group of those with <50 cells/mm^3, the hazard ratio (HR) was 0.56 (95% CI, 0.25–1.24) for those with a CD4 + T-cell count 51–100 cells/mm^3, 0.24 (95% CI, 0.08–0.68) for those with 101–200 cells/mm^3, and 0.15 (95% CI, 0.03–0.67) for those with >200 cells/mm^3. Age, weight, distance, WHO clinical stage and calendar year of ART initiation were not statistically significantly associated with death.

In the sample-corrected analysis (Table 2), deaths found through patient tracking were incorporated to represent outcomes in all lost patients. In multivariable analyses,

Table 2 Sample-corrected analysis of factors associated with mortality among ART initiators in Mbarara, Uganda. Results of unadjusted and multivariable proportional hazards regression models that incorporated deaths identified through tracking a representative sample of patients lost to follow-up in the community

	Unadjusted			Multivariable		
Factor	HR	*P*-value	95% CI	HR	*P*-value	95% CI
Age (per 10 years)	1.25	0.12	0.95–1.65	1.37	0.03	1.04–1.81
Male sex	1.11	0.37	0.65–1.92	1.02	0.96	0.57–1.83
Weight < 40 kg	3.63	<0.01	1.84–7.15	3.04	<0.01	1.58–5.85
Pre-ART CD4 count						
≤50 cells/mm^3	Ref			Ref		
51–100 cells/mm^3	0.58	0.22	0.24–1.37	0.58	0.21	0.24–1.36
101–200 cells/mm^3	0.20	<0.01	0.07–0.54	0.21	<0.01	0.08–0.56
>200 cells/mm^3	0.18	<0.01	0.05–0.66	0.20	0.02	0.06–0.73
Distance from home to clinic (per 10 km)	1.02	0.51	0.96–1.09	1.01	0.77	0.94–1.09
WHO clinical stage 4	1.27	0.48	0.64–2.52	0.82	0.54	0.43–1.55
Year of ART initiation						
2004	Ref			Ref		
2005	1.56	0.15	0.86–2.86	1.91	0.03	1.07–3.41
2006	0.62	0.27	0.26–1.46	0.88	0.78	0.37–2.08
2007	0.09	<0.01	0.02–0.41	0.15	0.02	0.03–0.71

ART, antiretroviral therapy; HR, hazard ratio.

pre-ART CD4 + T-cell count remained a significant predictor of death and the magnitude of association was similar to the naïve analysis. Male sex, however, which was found to be significantly associated with mortality in the naïve analysis, was not significantly associated with mortality in the sample-corrected analysis with a change in the magnitude of association from 1.86 to 1.02 (95% CI: 0.57–1.83). In contrast, age, which was not significantly associated with death in the naïve analysis, was significantly associated with death in the corrected analysis with each additional 10 years of age conferring a 37% increase in the rate of death (HR = 1.37, 95% CI: 1.04–1.81). Weight <40 kg at ART initiation (HR = 3.04, 95% CI: 1.58–5.85) and calendar year of ART initiation were also significantly associated with mortality in this analysis whereas in the naïve analysis they were not. In the sample-corrected analysis, there was a significant increase in the risk of death for subjects initiating ART in 2005 versus 2004 (the reference year) and a significant decrease in risk among those starting ART in 2007 compared to 2004.

A sensitivity analysis to explore the effect of residual bias from the 13% of patients in the tracked sample whose outcomes were not ascertained found no substantial differences. Neither a pessimistic model (which assumed all patients without vital status ascertainment had died) nor an optimistic model (which assumed they were all alive) changed the findings of multivariable model – all hazard ratios were similar and no *P*-values crossed the significance threshold.

Discussion

Understanding the magnitude and determinants of mortality in the millions of patients starting ART in resource-limited settings is needed to provide guidance for effective implementation of global ART delivery. Data obtained from routine care in clinic-based cohorts have been the first to emerge in the public health setting and to date has represented an important source of information (Ivers *et al.* 2005; Braitstein *et al.* 2006; Stringer *et al.* 2006; Toure *et al.* 2008). However, the high numbers of patients lost to follow-up in these settings threatens the validity of epidemiological analyses undertaken with them. These losses to follow-up represent one of the biggest barriers to global evidence-based delivery of ART.

We found that a sampling-based approach to identify determinants of mortality yielded major differences from the naïve approach and brings us closer to a true understanding of mortality on ART. First, the naïve analysis led to a spurious identification of male sex with death. Men appeared to have an 86% greater rate of death adjusted for other factors. Therefore, the nominal research implication would be to focus attention on men who start ART, and the nominal intervention would be to initiate intensive outreach efforts to men starting ART. A corrected analysis incorporating outcomes from a representative sample of patients lost to follow-up, however, eliminated this association. Policy in response to this spurious association would have lead to wasted resources, time and effort. Although our findings in southwestern Uganda are not necessarily generalizable across diverse socio-cultural settings across Africa, results from a number of existing studies in Africa that have found men to have higher rates of death (Ferradini *et al.* 2006; Stringer *et al.* 2006; Maskew 2008) should be considered in the light of the possible impact of losses to follow-up.

The naïve analysis also failed to identify three independent relationships between mortality age, calendar time of ART initiation and weight that were identified in the sample-corrected analysis. Age has been independently associated with death in African HIV-infected patients initiating ART in one study (Toure *et al.* 2008) but not in others (Ferradini *et al.* 2006; Johannessen *et al.* 2008). In industrialized countries, age has been consistently associated with blunted CD4 + T-cell recovery on ART (Florence *et al.* 2003; Kaufmann *et al.* 2005) and also found to predict mortality in patients on ART (Egger *et al.* 2002). Indeed, given the greater prevalence of opportunistic infections such as tuberculosis in Africa and the association between age and blunted immunologic recovery, the effect of age on mortality in Africa may be even more important than industrialized settings. Calendar year was another factor identified only by sample correction. Specifically, we found that patients starting ART in 2005 had the highest rate of death compared to the reference group of ART initiation in 2004, but then there was a decline in mortality in 2006 and 2007. During 2005, a large influx of patients initiated ART in the setting of expanded access. The clinic rapidly developed new protocols, acquired new staff and expanded physical capacity to accommodate the rapidly rising numbers of patients. As such, treatment outcomes may have improved after this initial 'programmatic learning curve' in the setting of rapid expansion. This period may have also included patients who were the sickest in the community and whose morbidity was not fully captured by the pre-ART CD4 + T-cell count, WHO clinical stage, or weight. Low pre-ART weight likely is associated with death because it represents advanced malnutrition or undiagnosed infections such as tuberculosis (Lucas *et al.* 1994). The role of weight has been consistently reported in other African studies (Lawn *et al.* 2006; Johannessen *et al.* 2008).

Although we are not the first to note that different methods of accommodating for the effects losses to follow-

up changes the outcomes of analyses, our approach extends previous work because it is conducted within a sampling framework that maximizes feasibility and scalability. Bisson *et al.* searched for all 68 patients lost to follow-up from a clinic in Botswana and found outcomes in 67.6%. They showed that accounting for the outcomes among the lost patients changed the association between death and both sex and pre-ART hemoglobin level. Given the vast scale of the global ART roll out and therefore the large numbers of patients lost to follow-up, a sampling-based approach is more feasible than searching for all lost patients and therefore potentially more scalable and sustainable. The analysis itself is also accessible and can be easily conducted on standard software (e.g. Stata, R, SAS, etc.). Finally, we believe ascertainment of vital status in a high proportion of a representative sample (87% in this case) likely leads to less biased findings than less complete ascertainment in a census-based approach.

Our sampling-based approach should be repeated in other ART delivery settings in resource-limited settings to strengthen the evidence for this approach. It is unknown whether outcomes in a representative sample can be identified in other African settings that differ by geography and culture and whether sample-corrected analyses will yield substantial departures from naive analyses in these settings. Furthermore, application in other settings may identify different associations with mortality: not all factors associated with death in southwestern Uganda can be generalized across the continent. In particular, the role of sex is likely modified by different socio-cultural environments and secular trends in mortality likely reflect local conditions. A sampling approach, conducted at the program level and with factors of particular local interest, can also enable programs to understand the most relevant determinants of mortality in their patients. Of note, this sampling-based approach is not meant to replace efforts to re-integrate as many lost patients into care as is possible. Instead, its goal is to strengthen programmatic efforts by providing valid data. Just as in clinical medicine where an accurate 'Assessment' is requisite for an appropriate 'Plan' public health strategies must be based on accurate epidemiologic assessments to be most effective.

There are limitations to this study. There are a number of potentially important predictor variables we did not have access to such as pre-ART hemoglobin and plasma HIV RNA level. Furthermore, because data on height were frequently missing, we used weight instead of body mass index (BMI) to represent energy deficiency. In addition, we did not find a significant association between WHO clinical stage and mortality in the corrected analysis. We attribute this to possible misclassification of WHO clinical stage, a common consequence of working in a setting with limited diagnostic capabilities. Indeed, recent data from the Infectious Diseases Institute in Kampala suggest that routine clinical diagnoses of WHO staging conditions are frequently misclassified when compared to a gold standard diagnostic approach (Kiragga *et al.* 2008). Finally, we note that this study seeks to identify valid determinants of mortality after correcting for losses to follow-up. There are, however, other important outcomes to understand, such as which patients cease to access care, which are not directly addressed by this analysis.

In summary, we have presented another application of a sampling-based approach to account for losses to follow-up in ART Programs in Africa. Not only can a sampling-based approach help derive accurate estimates of the absolute incidence of mortality after ART initiation, it can also be used to ascertain the determinants of mortality. Findings from a naïve analysis that did not account for losses to follow-up were dramatically biased, both in terms of (i) identification of spurious associations with mortality (male sex) and (ii) failure to identify significant associations (age, weight, calendar year of ART initiation). While we evaluated the consequences of losses to follow up on mortality, losses also have the potential to bias associations with other outcomes (i.e. CD4 + T-cell recovery, transfer, etc.). These biased findings have serious consequences: they can mislead public health interventions with costs in both dollars and more importantly potentially human lives. A sampling-based approach represents a potentially generalizable solution to the scientific obstacles created by losses to follow-up. This approach can strengthen the quality of evidence available for public health leadership and help to optimize global delivery of ART.

Acknowledgements

We are grateful to Hassan Baryahikwa, John Bennett, MPH, Este Hudes, PhD and Mark and Lisa Schwartz.

Conflicts of interest

The authors have declared that they have no conflicts of interest.

References

Braitstein P, Brinkhof MW, Dabis F *et al.* (2006) Mortality of HIV-1-infected patients in the first year of antiretroviral therapy: comparison between low-income and high-income countries. *Lancet* **367**, 817–824.

Egger M, May M, Chene G *et al.* (2002) Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a

collaborative analysis of prospective studies. *Lancet* **360**, 119–129.

Ferradini L, Jeannin A, Pinoges L *et al.* (2006) Scaling up of highly active antiretroviral therapy in a rural district of Malawi: an effectiveness assessment. *Lancet* **367**, 1335–1342.

Ferro-Luzzi A, Sette S, Franklin M & James WP (1992) A simplified approach of assessing adult chronic energy deficiency. *European Journal of Clinical Nutrition* **46**, 173–186.

Florence E, Lundgren J, Dreezen C *et al.* (2003) Factors associated with a reduced CD4 lymphocyte count response to HAART despite full viral suppression in the EuroSIDA study. *HIV Medicine* **4**, 255–262.

Frangakis CE & Rubin DB (2001) Addressing an idiosyncrasy in estimating survival curves using double sampling in the presence of self-selected right censoring. *Biometrics* **57**, 333–342.

Geng EH, Emenyonu N, Bwana MB, Glidden DV & Martin JN (2008) Sampling-based approach to determining outcomes of patients lost to follow-up in antiretroviral therapy scale-up programs in Africa. *JAMA* **300**, 506–507.

Geng EH, Bangsberg DR, Musinguzi N *et al.* (2009) Understanding Reasons for and Outcomes of Patients Lost to Follow-up in Antiretroviral Therapy Programs in Africa through a Sampling-based Approach. *JAIDS epub* **10**, 405–411.

Gilks CF, Crowley S, Ekpini R *et al.* (2006) The WHO public-health approach to antiretroviral treatment against HIV in resource-limited settings. *Lancet* **368**, 505–510.

Ivers LC, Kendrick D & Doucette K (2005) Efficacy of antiretroviral therapy programs in resource-poor settings: a meta-analysis of the published literature. *Clinical Infectious Diseases* **41**, 217–224.

Johannessen A, Naman E, Ngowi BJ *et al.* (2008) Predictors of mortality in HIV-infected patients starting antiretroviral therapy in a rural hospital in Tanzania. *BMC Infectious Diseases* **8**, 52.

Kaufmann GR, Furrer H, Ledergerber B *et al.* (2005) Characteristics, determinants, and clinical relevance of CD4 T cell recovery to <500 cells/microL in HIV type 1-infected individuals receiving potent antiretroviral therapy. *Clinical Infectious Diseases* **41**, 361–372.

Kiragga A, Castelnuovo B, Kambugu A, Schaefer P, Kamya M & Easterbrook P (2008) *Validation of the Quality of Data Collection in a Large Observational Clinic Database in Uganda Using a Nested Research Cohort*. International Workshop on HIV Observational Databases, Malaga.

Lawn SD, Myer L, Harling G, Orrell C, Bekker LG & Wood R (2006) Determinants of mortality and nondeath losses from an antiretroviral treatment service in South Africa: implications for program evaluation. *Clinical Infectious Diseases* **43**, 770–776.

Lin DY (2000) On fitting Cox's proportinal hazards models to survey data. *Biometrika* **87**, 37–47.

Lucas SB, De Cock KM, Hounnou A *et al.* (1994) Contribution of tuberculosis to slim disease in Africa. *BMJ* **308**, 1531–1533.

Maskew M (2008) *Gender Differences in HIV Outcomes in South African Patients Receiving Highly Active Antiretroviral Therapy*. 2008 HIV/AIDS Implementer's Meeting, Kampala.

Rosen S, Fox MP & Gill CJ (2007) Patient retention in antiretroviral therapy programs in sub-Saharan Africa: a systematic review. *PLoS Medicine* **4**, e298.

Rothman KJ & Greenland S (2008) *Modern Epidemiology*. Lippincott, Philadelphia.

Schafer JL (1999) Multiple imputation: a primer. *Statistical Methods in Medical Research* **8**, 3–15.

Stringer JS, Zulu I, Levy J *et al.* (2006) Rapid scale-up of antiretroviral therapy at primary care sites in Zambia: feasibility and early outcomes. *JAMA* **296**, 782–793.

Toure S, Kouadio B, Seyler C *et al.* (2008) Rapid scaling-up of antiretroviral therapy in 10,000 adults in Cote d'Ivoire: 2-year outcomes and determinants. *AIDS* **22**, 873–882.

Wools-Kaloustian K, Kimaiyo S, Diero L *et al.* (2006) Viability and effectiveness of large-scale HIV treatment initiatives in sub-Saharan Africa: experience from western Kenya. *AIDS* **20**, 41–48.

Yiannoutsos CT, An MW, Frangakis CE *et al.* (2008) Sampling-based approaches to improve estimation of mortality among patient dropouts: experience from a large PEPFAR-funded program in Western Kenya. *PLoS ONE* **3**, e3843.

Corresponding Author Elvin H. Geng, Division of HIV/AIDS at San Francisco General Hospital, University of California San Francisco, 995 Potrero Avenue, Building 80, San Francisco, CA 94110, USA. Tel.: +1 415 430 5589; Fax +1 415 476 6953; E-mail: genge@php.ucsf.edu