Practice of Epidemiology

An Empirical Approach to Defining Loss to Follow-up Among Patients Enrolled in Antiretroviral Treatment Programs

Benjamin H. Chi*, Ronald A. Cantrell, Albert Mwango, Andrew O. Westfall, Wilbroad Mutale, Mohammed Limbada, Lloyd B. Mulenga, Sten H. Vermund, and Jeffrey S. A. Stringer

* Correspondence to Dr. Benjamin H. Chi, Centre for Infectious Disease Research in Zambia, Box 34681, 1275 Lubuto Road, Lusaka, Zambia (e-mail: bchi@cidrz.org).

Initially submitted October 5, 2009; accepted for publication January 8, 2010.

In many programs providing antiretroviral therapy (ART), clinicians report substantial patient attrition; however, there are no consensus criteria for defining patient loss to follow-up (LTFU). Data on a multisite human immunodeficiency virus (HIV) treatment cohort in Lusaka, Zambia, were used to determine an empirical “days-late” definition of LTFU among patients on ART. Cohort members were classified as either “in care” or LTFU as of December 31, 2007, according to a range of days-late intervals. The authors then looked forward in the database to determine which patients actually returned to care at any point over the following year. The interval that best minimized LTFU misclassification was described as “best-performing.” Overall, 33,704 HIV-infected adults on ART were included. Nearly one-third (n = 10,196) were at least 1 day late for an appointment. The best-performing LTFU definition was 56 days after a missed visit, which had a sensitivity of 84.1% (95% confidence interval (CI): 83.2, 85.0), specificity of 97.5% (95% CI: 97.3, 97.7), and misclassification of 5.1% (95% CI: 4.8, 5.3). The 60-day threshold performed similarly well, with only a marginal difference (<0.1%) in misclassification. This analysis suggests that >60 days since the last appointment is a reasonable definition of LTFU. Standardization to empirically derived definitions of LTFU will permit more reliable comparisons within and across programs.

Africa; antiretroviral therapy, highly active; follow-up studies; HIV; patient dropouts; Zambia

Abbreviations: ART, antiretroviral therapy; CI, confidence interval; HIV, human immunodeficiency virus.
phenomenon of patient attrition, there is a need for standardized definitions of loss to follow-up based on empirical evidence, to permit consistent comparisons across and within programs.

MATERIALS AND METHODS

We analyzed data from a large programmatic cohort in Lusaka, Zambia, to determine the best-performing criterion for classifying patients who are late for scheduled appointments. The multisite Lusaka ART program is administered by the Zambian Ministry of Health and its local partners, with substantial support being provided by the President’s Emergency Plan for AIDS Relief; the Global Fund for AIDS, Tuberculosis, and Malaria; and other donors. Clinical and programmatic characteristics have been described in detail elsewhere (4, 5). Briefly, HIV-infected adults and children are screened for ART eligibility on the basis of CD4+ cell count and clinical staging across 18 public sector sites. ART is initiated on the basis of national criteria (17), which closely follow those of the World Health Organization (18, 19). Adults initiating ART undergo an intensive clinical visit schedule over the first 3 months for assessment of potential side effects and encouragement of adherence; patients without complications proceed to once-monthly pharmacy visits and quarterly clinical visits. Medical history and appointment information is captured in an electronic medical record (20). From this database, data staff generate weekly lists of patients with missed appointments. Community health workers use collected locator information to contact patients with missed visits at their homes to encourage and facilitate clinic attendance (21).

We developed an empirical approach to determine the “most efficient” definition of loss to follow-up (Figure 1) based on sensitivity and specificity to predict a return to care in the subsequent year. The source population for our analysis was all ART-naive HIV-infected adults initiating HIV treatment at 18 Lusaka sites between April 1, 2004, and December 31, 2007. Patients who were active and those who were late for a clinical or pharmacy visit as of December 31, 2007, were included in the analysis. Patients who had formally withdrawn from the program or who had died prior to that date were excluded (22). We categorized loss to follow-up as of December 31, 2007, but failed to return for subsequent appointments (false negatives). Group d comprises patients who were classified as active at the cutoff date and remain active during the coming year (true negatives). Using this 2 × 2 table, we are able to calculate sensitivity (a/a + c) and specificity (d/b + d) for each LTFU threshold.
characteristic curves. “False-positive” cases were defined as persons who were classified as follow-up losses on December 31, 2007, but returned to care in the subsequent year. The false-positive rate for each cutoff definition of loss to follow-up (LTFU) was calculated as (non-LTFU prevalence) \( \times (1 - \text{specificity}) \). “False-negative” cases for each cutoff definition of loss to follow-up were persons who were classified as active but never returned for later visits; this was defined as (LTFU prevalence) \( \times (1 - \text{sensitivity}) \). The cutpoint that minimized the sum of false positives and false negatives was considered the most efficient loss to follow-up threshold—that is, a weighted sum of the sensitivity and specificity based on the prevalence of cases (23). If 2 or more time intervals had the same misclassification rate, we designated the shorter interval as the more efficient definition of loss to follow-up.

To determine whether length of enrollment affected the performance of our calculated definitions of loss to follow-up, we performed secondary analyses using different “enrollment cohorts.” We stratified patients according to the calendar year in which they started ART and, using the same method, calculated the optimal days-late threshold for defining loss to follow-up. These cohorts comprised persons starting ART between: 1) April 1, 2004, and December 31, 2004; 2) January 1, 2005, and December 31, 2005; 3) January 1, 2006, and December 31, 2006; and 4) January 1, 2007, and December 31, 2007. For each of these cohorts, December 31, 2007, was used as the cutpoint to classify patients as lost to follow-up, and we looked ahead 1 year to determine their subsequent status.

Two additional secondary analyses were performed. First, we restricted our study population to persons who were at least 1 day late for their clinical or pharmacy appointment as of December 31, 2007. We excluded the subset of patients who were active in care and had their next appointment following this cutoff date, because they could only contribute to “false-negative” misclassification. In contrast, those included in this subanalysis could potentially contribute to both “false-positive” and “false-negative” misclassification, depending on the definition of loss to follow-up used. Secondly, we tested the performance of definitions of loss to follow-up at other cutoff dates. For the primary analysis, December 31, 2007, had been chosen as a matter of convenience, but an underlying tenet of this method is that any date can be used, provided that a sufficient amount of follow-up time is available afterwards. We conducted this secondary analysis to confirm the robustness of our findings when other dates were used.

All analyses were performed using SAS, version 9.13 (SAS Institute Inc., Cary, North Carolina). Use of these programmatic data was approved by the University of Zambia (Lusaka, Zambia) Research Ethics Committee and the
University of Alabama at Birmingham (Birmingham, Alabama) Institutional Review Board.

RESULTS

Between April 1, 2004, and December 31, 2007, 40,700 HIV-infected adults initiated ART at 18 program sites in Lusaka. As of December 31, 2007, 6,996 (17.2%) were known to have formally withdrawn from the program or to have died and were thus excluded. Of the remaining 33,704 patients, 23,508 (69.7%) had attended their last appointment and had a subsequent one scheduled and were thus considered “active.” Nearly one-third of patients (n = 10,196 or 30.3%) were at least 1 day late for a clinical or pharmacy visit. The median interval from the last missed appointment to the December 31, 2007, eligibility cutpoint was 123 days (interquartile range, 21–462). The distributions of the yearly enrollment cohorts are shown in Figure 2.

We evaluated the performance of various cutpoints for defining loss to follow-up, including sensitivity, specificity, and misclassification rate (Table 1), and then fitted receiver operating characteristic curves to these data (Figure 3). On the basis of our method, the best-performing definition was \( \geq 56 \) days. This threshold had the lowest misclassification rate (5.1%; 95% confidence interval (CI): 4.8, 5.3), along with a sensitivity of 84.1% (95% CI: 83.2, 85.0), specificity of 97.5% (95% CI: 97.3, 97.7), a positive predictive value of 88.7% (95% CI: 87.9, 89.5), and a negative predictive value of 96.3% (95% CI: 96.1, 96.5). Interestingly, misclassification rates for \( \geq 54 \) days to \( \geq 69 \) days differed by less than 0.1% when compared with the misclassification rate of \( \geq 56 \) days (Figure 4, part A).

When we performed analyses of separate enrollment cohorts, the most efficient definition of loss to follow-up appeared to shorten as duration of time in the program decreased (Table 2). For persons enrolled for 3 years or longer (i.e., those who started ART prior to December 31, 2004), we found that the best-performing definition was \( \geq 42 \) days, which had a sensitivity of 86.4% (95% CI: 85.5, 87.4), specificity of 96.6% (95% CI: 96.5, 96.6), a positive predictive value of 89.9% (95% CI: 89.6, 90.2), and a negative predictive value of 97.3% (95% CI: 97.2, 97.4). These rates were similar to those for persons enrolled for 2 years or longer (i.e., those who started ART prior to December 31, 2005), who showed that the best-performing definition was \( \geq 49 \) days, which had a sensitivity of 85.4% (95% CI: 84.5, 86.3), specificity of 97.1% (95% CI: 97.0, 97.2), a positive predictive value of 88.7% (95% CI: 88.0, 89.4), and a negative predictive value of 96.3% (95% CI: 96.1, 96.5). Interestingly, misclassification rates for \( \geq 47 \) days to \( \geq 53 \) days differed by less than 0.1% when compared with the misclassification rate of \( \geq 49 \) days (Figure 4, part B).

#### Table 1. Performance of Various Time Intervals After a Missed Appointment in Defining Loss to Follow-up Among Patients Enrolled in a Multisite Human Immunodeficiency Virus Treatment Cohort Between April 1, 2004, and December 31, 2007, Lusaka, Zambia

<table>
<thead>
<tr>
<th>LTFU Threshold(^a), days</th>
<th>No. Classified as LTFU</th>
<th>No. Who Subsequently Returned to Care</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>Misclassification, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \geq 7 )</td>
<td>9,269</td>
<td>3,308</td>
<td>92.6</td>
<td>87.9</td>
<td>11.2</td>
</tr>
<tr>
<td>( \geq 14 )</td>
<td>8,397</td>
<td>2,539</td>
<td>91.0</td>
<td>90.7</td>
<td>9.3</td>
</tr>
<tr>
<td>( \geq 21 )</td>
<td>7,686</td>
<td>1,907</td>
<td>89.8</td>
<td>93.0</td>
<td>7.6</td>
</tr>
<tr>
<td>( \geq 28 )</td>
<td>7,063</td>
<td>1,377</td>
<td>88.3</td>
<td>95.0</td>
<td>6.3</td>
</tr>
<tr>
<td>( \geq 35 )</td>
<td>6,730</td>
<td>1,110</td>
<td>87.3</td>
<td>95.9</td>
<td>5.7</td>
</tr>
<tr>
<td>( \geq 42 )</td>
<td>6,499</td>
<td>938</td>
<td>86.4</td>
<td>96.6</td>
<td>5.4</td>
</tr>
<tr>
<td>( \geq 49 )</td>
<td>6,295</td>
<td>798</td>
<td>85.4</td>
<td>97.1</td>
<td>5.2</td>
</tr>
<tr>
<td>( \geq 56 )(^b)</td>
<td>6,104</td>
<td>688</td>
<td>84.1</td>
<td>97.5</td>
<td>5.1</td>
</tr>
<tr>
<td>( \geq 63 )</td>
<td>5,958</td>
<td>617</td>
<td>83.0</td>
<td>97.7</td>
<td>5.1</td>
</tr>
<tr>
<td>( \geq 70 )</td>
<td>5,841</td>
<td>576</td>
<td>82.8</td>
<td>97.9</td>
<td>5.2</td>
</tr>
<tr>
<td>( \geq 77 )</td>
<td>5,714</td>
<td>528</td>
<td>80.6</td>
<td>98.1</td>
<td>5.3</td>
</tr>
<tr>
<td>( \geq 84 )</td>
<td>5,603</td>
<td>493</td>
<td>79.4</td>
<td>98.2</td>
<td>5.4</td>
</tr>
<tr>
<td>( \geq 91 )</td>
<td>5,514</td>
<td>462</td>
<td>78.5</td>
<td>98.3</td>
<td>5.5</td>
</tr>
<tr>
<td>( \geq 98 )</td>
<td>5,431</td>
<td>441</td>
<td>77.5</td>
<td>98.4</td>
<td>5.6</td>
</tr>
<tr>
<td>( \geq 105 )</td>
<td>5,333</td>
<td>421</td>
<td>76.3</td>
<td>98.5</td>
<td>5.8</td>
</tr>
<tr>
<td>( \geq 112 )</td>
<td>5,235</td>
<td>403</td>
<td>75.1</td>
<td>98.5</td>
<td>6.0</td>
</tr>
<tr>
<td>( \geq 119 )</td>
<td>5,147</td>
<td>383</td>
<td>74.0</td>
<td>98.6</td>
<td>6.1</td>
</tr>
<tr>
<td>( \geq 126 )</td>
<td>5,073</td>
<td>371</td>
<td>73.1</td>
<td>98.6</td>
<td>6.3</td>
</tr>
<tr>
<td>( \geq 133 )</td>
<td>4,991</td>
<td>356</td>
<td>72.0</td>
<td>98.7</td>
<td>6.4</td>
</tr>
<tr>
<td>( \geq 140 )</td>
<td>4,910</td>
<td>344</td>
<td>70.9</td>
<td>98.7</td>
<td>6.6</td>
</tr>
<tr>
<td>( \geq 147 )</td>
<td>4,837</td>
<td>329</td>
<td>70.0</td>
<td>98.8</td>
<td>6.7</td>
</tr>
<tr>
<td>( \geq 154 )</td>
<td>4,780</td>
<td>317</td>
<td>69.3</td>
<td>98.8</td>
<td>6.8</td>
</tr>
<tr>
<td>( \geq 161 )</td>
<td>4,724</td>
<td>309</td>
<td>68.6</td>
<td>98.9</td>
<td>6.9</td>
</tr>
<tr>
<td>( \geq 168 )</td>
<td>4,656</td>
<td>297</td>
<td>67.7</td>
<td>98.9</td>
<td>7.1</td>
</tr>
<tr>
<td>( \geq 175 )</td>
<td>4,581</td>
<td>286</td>
<td>66.7</td>
<td>99.0</td>
<td>7.2</td>
</tr>
<tr>
<td>( \geq 182 )</td>
<td>4,511</td>
<td>275</td>
<td>65.8</td>
<td>99.0</td>
<td>7.4</td>
</tr>
</tbody>
</table>

Abbreviation: LTFU, loss to follow-up.

\(^a\) For illustrative purposes, we show the performance of different LTFU time periods in 7-day increments.

\(^b\) The best-performing definition of LTFU coincided with the 56-day intervals.
2004), the most efficient definition of follow-up loss was 70 days. For persons enrolled for at least 2 years but less than 3 years (i.e., January 1, 2005–December 31, 2005), the most efficient definition of follow-up loss was 63 days. For persons enrolled for at least 1 year but less than 2 years (i.e., January 1, 2006–December 31, 2006), the most efficient definition of follow-up loss was 60 days. For persons enrolled for less than 1 year (i.e., January 1, 2007–December 31, 2007), the most efficient definition of follow-up loss was 47 days.

We restricted the study population to persons who were at least 1 day late for their last clinical or pharmacy appointment as of December 31, 2007. When we evaluated the performance of loss to follow-up classifications at different time intervals, the best-performing definition was ≥56 days, identical to our primary analysis (Figure 4, part B). The proportion of patients misclassified at this threshold was 12.8% (95% CI: 12.1, 13.4); sensitivity was 89.8% (95% CI: 89.0, 90.6), specificity was 83.5% (95% CI: 82.3, 84.6), positive predictive value was 88.7% (95% CI: 87.9, 89.5), and negative predictive value was 85.0% (95% CI: 83.6, 86.1). We then evaluated the performance of definitions of loss to follow-up at cutoff dates other than December 31, 2007. This analysis yielded results nearly identical to those of our primary analysis (Figure 4, part C). The most efficient definitions using alternative cutoff days were: June 30, 2007 (62 days); September 30, 2007 (60 days); March 31, 2008 (62 days); and June 30, 2008 (59 days).

DISCUSSION

Minimizing the misclassification of patients lost to follow-up has great importance in cohort analyses and programmatic reporting. When patients are prematurely classified as lost to follow-up, they can be incorrectly censored and fail to contribute person-time to an analysis, thus contributing to underestimation of program coverage over time.

Figure 3. Receiver operator characteristic curves for adults initiating antiretroviral therapy in Lusaka, Zambia, between April 1, 2004, and December 31, 2007, by year of enrollment.

Figure 4. Proportions of patients misclassified as active or lost to follow-up across different time intervals among adults initiating antiretroviral therapy in Lusaka, Zambia, between April 1, 2004, and December 31, 2007. In a series of sensitivity analyses, we stratified the overall population according to year of enrollment (part A); restricted our study population to only those patients who were at least 1 day late for a scheduled visit on December 31, 2007 (part B); and repeated our analysis using several cutoff dates other than December 31, 2007 (part C).
For this reason, many cohort studies have utilized more liberal definitions of loss to follow-up to maximize the possibility of return. When this window is too wide, however, patients may be misclassified as active even when they will not return. A large proportion may in fact have died (14, 21, 24, 25). Here we have proposed a simple empirical approach to defining loss to follow-up in a way that can standardize program evaluation reports and research comparisons, both across and within programs. In our multisite program in Lusaka, use of ≥56 days’ lateness for a scheduled visit led to the fewest misclassifications of loss to follow-up. However, given the minute differences in misclassification seen in the intervals immediately preceding and following 56 days (Figure 4), we suggest the use of 2 months—or 60 days—as the lateness threshold for defining loss to follow-up. This interval is longer than the definition of 30 days that we have used in our previous program evaluations (4–11) but shorter than that used in many other programs (2, 14–16).

The strengths of this method are its relative simplicity and empirical approach, suggesting a standardized, evidence-based definition of loss to follow-up. We recognize several limitations to our analysis. This technique only considers the patient’s last missed visit; we did not take into account previous missed visits and thus failed to utilize all available data. Our approach was designed to most accurately classify patients as either active or lost to follow-up at a single point in time, as would be required for program reporting or cohort analysis. This method is not conducive to more complicated analysis regarding factors associated with loss to follow-up, since persons who died or withdrew from the program prior to the date of classification were excluded. Although our definition of follow-up loss was robust across a number of sensitivity analyses, it is unknown whether it can be generalized to locations outside of our urban primary-care setting in Zambia. Several program characteristics may have affected our results, including an active patient tracking system for missed visits, the provision of free ART, and the use of sophisticated electronic medical records linking care data across multiple clinics. Similar analyses should be performed in other settings to determine how these and other program features influence optimal definitions of loss to follow-up.

Our method begins with the selection of an arbitrary date (i.e., December 31, 2007), from which patients are categorized as active or lost to follow-up according to a range of days-late thresholds. Once categorized as lost to follow-up, all patients are provided the same opportunity to return to care irrespective of time since the last visit; in this analysis, this was the calendar year of January 1, 2008–December 31, 2008. We recognize, however, that the likelihood that such a person will return to care varies inversely with the length of time since his or her missed visit. A patient whose last scheduled visit was 2 years ago, for example, is far less likely to return to care than someone who missed a visit just 2 months ago. Because the chances of such misclassification may be reduced among “older” enrollment cohorts, the performance of the “most efficient” definition of loss to follow-up was generally better in those groups. While specificity remained approximately the same, sensitivity increased among persons who had been enrolled in the program for a longer time (Table 2).

When we examined the proportion of ART patients misclassified as either active or lost to follow-up, the greatest change in misclassification was seen over the first 30 days. There appeared to be a broad nadir in the 30- to 90-day range, followed by a slight increase over the intervals that followed. While our method can determine the best-performing definition of loss to follow-up with great precision, as Figure 4 demonstrates, there probably exists a range of “acceptable” thresholds from the program perspective. This flexibility prompted us to recommend a 60-day threshold for defining loss to follow-up, which we found to be more intuitive and only marginally less efficient than the best-performing 56-day threshold in our setting.

We observed differences in the optimal definition of loss to follow-up when stratifying by enrollment date. Persons who had enrolled most recently had the shortest loss to follow-up windows; as time in the program increased, so did the length of this best-performing interval. The reasons behind this finding are unclear, but our exclusion of deaths and program withdrawals from the analysis may have played an important role. If 1 group were to have better status ascertainment—a likely scenario for earlier (vs. later) enrollment cohorts—then the differential exclusion of these patients could contribute to the variability from stratum to stratum.

When we investigated the interval between the missed visit and December 31, 2007, we found a bimodal distribution of these data for each enrollment cohort. A common
feature of each was the prominent peak within 90 days, suggesting that the majority of missed visits had occurred only a few months earlier. However, we consistently observed a smaller peak for follow-up losses, generally coinciding with the year of enrollment for each cohort. This suggests a high rate of attrition early in a patient’s ART course, with unreported early mortality being a likely and important contributor to these losses (26).

Arbitrary and variable definitions of loss to follow-up in ART programs are now extant, limiting the usefulness of monitoring and evaluation activities within and across programs. Here we have proposed a simple approach to determining the best-performing criterion for classifying patients as lost to follow-up. On the basis of data from the multisite Lusaka program, we suggest a 2-month (60-day) lateness threshold for defining loss to follow-up. Since the basic data needed to calculate such thresholds are found in many electronic medical records in sub-Saharan Africa (27), this method should be considered for program-, country-, and/or region-specific definitions of loss to follow-up.

ACKNOWLEDGMENTS

Author affiliations: Centre for Infectious Disease Research in Zambia, Lusaka, Zambia (Benjamin H. Chi, Ronald A. Cantrell, Andrew O. Westfall, Mohammed Limbada, Lloyd B. Mulenga, Sten H. Vermund, Jeffrey S. A. Stringer); Department of Obstetrics and Gynecology, School of Medicine, University of Alabama at Birmingham, Birmingham, Alabama (Benjamin H. Chi, Ronald A. Cantrell, Andrew O. Westfall, Jeffrey S. A. Stringer); Zambian Ministry of Health, Lusaka, Zambia (Albert Mwango); Department of Community Medicine, School of Medicine, University of Zambia, Lusaka, Zambia (Wilbroad Mubana); and Institute for Global Health, School of Medicine, Vanderbilt University, Nashville, Tennessee (Sten H. Vermund).

This work was supported in part by the President’s Emergency Plan for AIDS Relief through a multicountry grant to the Elizabeth Glaser Pediatric AIDS Foundation from the US Department of Health and Human Services and the US Centers for Disease Control and Prevention’s Global AIDS Program (cooperative agreement U62/CCU12354). Additional investigator salary or trainee support was provided by the US National Institutes of Health (grants K01-TW06670, D43-TW001035, and P30-AI027767) and a Clinical Scientist Development Award from the Doris Duke Charitable Foundation (grant 2007061).

The authors acknowledge the Zambian Ministry of Health for consistent and high-level support of operations research in the context of HIV program expansion.

The findings and conclusions included herein are solely the responsibility of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention. The funding agencies played no role in study design, data collection, data analysis, or manuscript writing.

Conflict of interest: none declared.

REFERENCES


