A Two-stage Approach for Rapid Assessment of the Proportion Achieving Viral Suppression Using Routine Clinical Data

Jessie K. Edwards,[a] *Yeycy Donastorg,*[b] *Sabrina Zadrozny,*[c] *Sarah Hileman,*[d] *Hoisex Gómez,*[b] *Marissa J. Seamans,*[e] *Michael E. Herce,*[f] *Edwin Ramírez,*[g] *Clare Barrington,*[h,i] *and Sharon Weir*[a,i]

Background: Improving viral suppression among people with HIV reduces morbidity, mortality, and transmission. Accordingly, monitoring the proportion of patients with a suppressed viral load is important to optimizing HIV care and treatment programs. But viral load data are often incomplete in clinical records. We illustrate a two-stage approach to estimate the proportion of treated people with HIV who have a suppressed viral load in the Dominican Republic.

Methods: Routinely collected data on viral load and patient characteristics were recorded in a national database, but 74% of patients on treatment at the time of the study did not have a recent viral load measurement. We recruited a subset of these patients for a rapid assessment that obtained additional viral load measurements. We combined results from the rapid assessment and main database using a two-stage weighting approach and compared results to estimates obtained using standard approaches to account for missing data.

Results: Of patients with recent routinely collected viral load data, 60% had a suppressed viral load. Results were similar after applying standard approaches to account for missing data. Using the two-stage approach, we estimated that 77% (95% confidence interval [CI] = 74, 80) of those on treatment had a suppressed viral load.

Conclusions: When assessing the proportion of people on treatment with a suppressed viral load using routinely collected data, applying standard approaches to handle missing data may be inadequate. In these settings, augmenting routinely collected data with data collected through sampling-based approaches could allow more accurate and efficient monitoring of HIV treatment program effectiveness.

Keywords: Epidemiologic biases; HIV; Routinely collected health data; Viral load

(*Epidemiology* 2022;33: 642–649)

Submitted November 9, 2021; accepted May 25, 2022

From the [a]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC; [b]Instituto Dermatológico y Cirugia de Piel, Santo Domingo, Dominican Republic; [c]Frank Porter Graham Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC; [d]Independent, Chapel Hill, NC; [e]Department of Epidemiology, University of Los Angeles, Los Angeles, CA; [f]School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC; [g]Servicio Nacional de Salud, Santo Domingo, Dominican Republic; [h]Department of Health Behavior, University of North Carolina at Chapel Hill, Chapel Hill, NC; and [i]Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC.

The research upon which this publication was based was funded by the generous support of the American people through the United States Agency for International Development (USAID) and the U.S. President's Emergency Plan for AIDS Relief (PEPFAR) through the Linkages across the Continuum of HIV Services for Key Populations Affected by HIV (LINKAGES) Project. LINKAGES, a 5-year cooperative agreement (AID-OAA-A-14-00045), is the largest global project dedicated to key populations. LINKAGES is led by FHI 360 in partnership with IntraHealth International, Pact, and the University of North Carolina at Chapel Hill. Data analysis and article preparation were supported in part by the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Additional funding for article preparation was provided by the National Institutes of Health through K01AI125087 and R01AI57758.

The authors report no conflicts of interest.

The contents of this publication do not necessarily reflect the views of National Institutes of Health, United States Agency for International Development, U.S. President's Emergency Plan for AIDS Relief, or the United States Government.

To protect patient confidentiality, data and code for replication are not available for this study.

Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Correspondence: Jessie K. Edwards, University of North Carolina at Chapel Hill, 2101 McGavran-Greenberg Hall, CB #7435, Chapel Hill, NC 27599. E-mail: jessedwards@unc.edu.

ISSN: 1044-3983/22/335-642-649
DOI: 10.1097/EDE.0000000000001513

A primary goal of HIV care and treatment programs is to maximize the amount of time patients with HIV spend with a suppressed viral load.[1–6] Viral load is an important clinical endpoint; viral suppression reduces onward transmission and improves the long-term health and quality of life for people living with HIV.[7] Accordingly, viral suppression is a key component of the Joint United Nations Programme on HIV/AIDS (UNAIDS) 95–95–95 goals, which state that 95% of people living with HIV should know their status, 95% of people who know their status should be treated with combination antiretroviral therapy (ART), and 95% of treated patients should achieve viral suppression.[8]

The proportion of people with HIV with a suppressed viral load is an important measure to inform HIV prevention, care, and treatment programs at the national, regional, and community levels. This measure has often been estimated using large

population-based studies.[9–11] However, such studies are resource intensive and may be difficult to conduct with the frequency required to measure and act on progress towards the 95–95–95 targets over time. In the absence of information from a population-based survey, information on viral load from routine clinical data may be leveraged to monitor progress towards the "third 95" target.

However, in many resource-constrained settings, people living with HIV do not have frequent viral load measurements even while on treatment and retained in care, resulting in substantial missing viral load data in clinical databases.[12] Even in settings recommending routine viral load assessment at set intervals, viral load testing may be used by some providers as a diagnostic tool (i.e., to confirm virologic failure in cases of clinical deterioration) rather than as a tool for monitoring viral suppression. In addition, routine viral load testing may be more likely to be conducted for patients who have access to private clinics or display strong health seeking behaviors than for other patients. The extensive amount of missing viral load data in many clinical settings means that the logical bounds on the proportion of people on treatment with a suppressed viral load are wide and uninformative. In addition, because people with and without routinely collected viral load data are likely different, but factors affecting viral load testing are not captured in routine care and treatment databases, standard approaches to account for missing data[13] often will not provide consistent estimates of the proportion of patients with a suppressed viral load.

We propose an approach to estimate the proportion of patients in HIV care and on treatment with a suppressed viral load using two-stage[14] (or double sampling[15,16]) methods. This approach augments routinely collected viral load data (which are typically collected not at random) with additional viral load measurements from a sample of those missing routinely collected viral load data. We illustrate this approach to estimate the proportion of people treated for HIV who had a suppressed viral load in the Dominican Republic. We compare estimates of the proportion with a suppressed viral load using only routinely collected viral load measurements, after applying standard statistical approaches to account for missing data, and after implementation of the two-stage approach.

METHODS

The parameter of interest was the cross-sectional proportion of people in care for HIV on antiretroviral therapy (ART) with a suppressed viral load on a specific target date (25 January 2017).

Outcome Definitions

Patients were considered to be "on ART" if they had started ART before the target date and were "retained in care" on this date, which we defined as having had at least one clinic visit within the 6 months before the target date. Viral suppression was defined as at least one viral load measurement in the 6 months before the target date below 200 copies/mL.[17] If a patient had more than one viral load measurement in the 6 months before the target date, we used the measurement closest to the target date. We considered viral loads to be missing for patients who were on ART and engaged in care without a viral load measurement in the 6 months before the target date.

First Stage: Routinely Collected Data From a National Database

Clinical data on the population of people living with HIV in the Dominican Republic are captured in routine clinical records and uploaded into a national database. At the time of the study, the national database included over 40,000 patients in care at 72 HIV care and treatment clinics in the country. At a patient's first visit to any HIV care and treatment clinic, baseline information on demographics, medications dispensed, and health status are recorded. An additional record is created for each follow-up visit to record treatments received and biomarker values. Information on date of death is included in the national database through linkage to the national vital statistics registry. At the time of this analysis, the recommended frequency for viral load monitoring was every 6 months.

For this analysis, we included all patients (a) who entered HIV care and treatment in the Dominican Republic between roll-out of the national database system on 1 June 2013 and 25 January 2017; (b) had started ART before 25 January 2017; and (c) were retained in care on 25 January 2017.

Second Stage: Rapid Assessment

We performed a rapid assessment to estimate and compare the probability of viral suppression among patients without a routinely collected viral load measurement. We then used this information to estimate the overall proportion suppressed by combining data from the rapid assessment with the routinely collected viral load data.

The rapid assessment was performed between 25 January 2017 and 10 March 2017 at four non-governmental organization (NGO) and five public HIV care and treatment facilities across the Dominican Republic and consisted of viral load measurements and a short patient survey. At the time of the study, approximately 25% of patients in care for HIV received care at a site selected for the assessment. Patients were eligible for the assessment if they entered HIV care and treatment in the Dominican Republic between 1 June 2013 and 25 January 2017 and were age 18 or older. During the rapid assessment, we approached an unselected consecutive sample of 1,084 patients arriving for a routine HIV care and treatment visit at any of the nine study sites. Among the 1,047 patients who provided written informed consent to participate (97% of those approached), facility staff administered a short survey about health and health behaviors, examined clinical records, and asked participants to provide a sample of whole blood for viral load testing. Participants were not required to provide a blood sample to participate in the assessment; 92 patients participated in the survey without providing a blood sample.

Of the 955 patients who provided a blood sample for a viral load measurement during the assessment, 452 had started ART, were retained in care on 25 January 2017, and had not had a routine viral load measurement within the past 6 months in the national database. This subset of 452 patients composed a supplemental sample used in later steps of the analysis (the additional 503 who participated in the assessment had nonmissing viral load data in the national database and therefore were already represented in the study data). All participants in the rapid assessment provided written informed consent. This project was approved by the Comisión Nacional de Bioética en Salud in the Dominican Republic and the Institutional Review Board at the University of North Carolina at Chapel Hill.

Statistical Methods

We compared estimates of the proportion with a suppressed viral load: (1) under a complete-case analysis based on the national database alone; (2) using logical bounds under the observed amount of missing data; (3) using standard statistical methods for missing data (multiple imputation and inverse probability weighting); and (4) applying the proposed two-stage approach. For the complete-case analysis, we restricted the sample to patients on ART and retained in care with a viral load measurement recorded in the national database in the past 6 months. The estimated proportion suppressed was the number with a suppressed viral load divided by the number of those on ART and retained in care with any viral load measurement during the relevant time period. This proportion would be expected to provide an accurate estimate of viral suppression among patients in care and on ART if viral load information were missing completely at random in this group.[13]

Next, to examine the uncertainty in this estimated proportion due to missing data, we calculated the logical bounds on the proportion suppressed under two extreme assumptions about missing data: for the upper bound, we assumed that all patients with missing viral load measurements had a suppressed viral load, and for the lower bound, we assumed that all patients with missing viral load measurements had an unsuppressed viral load. These proportions provide the range of estimates compatible with the observed data.

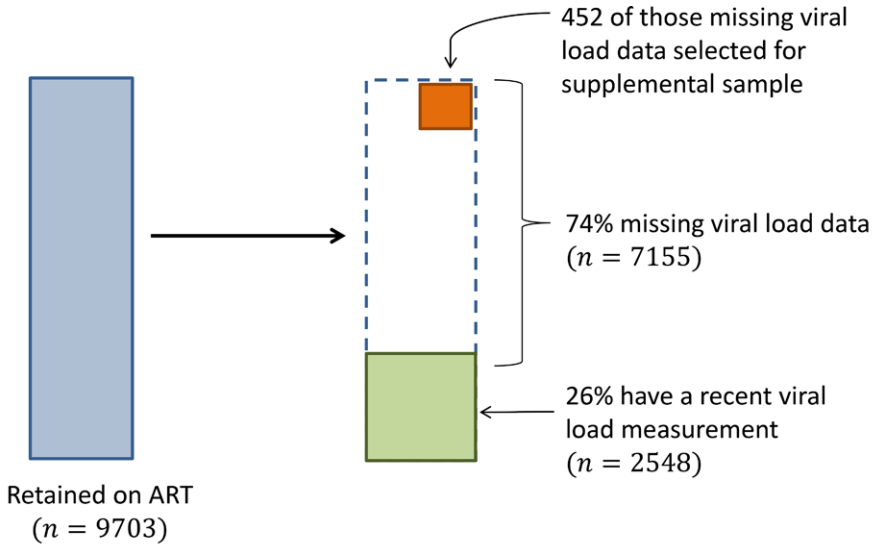
Third, we used two standard statistical methods for handling missing data to relax the assumption that viral loads were missing completely at random: (a) multiple imputation and (b) inverse probability weighting. Using multiple imputation,[18] we imputed viral loads for patients with missing viral load data based on demographic characteristics and other data recorded in the national database. To do this, in the subset of patients with a recent viral load measurement recorded in the database, we fit a logistic regression model for viral suppression conditional on measured covariates age, sex, year of entry into care, nationality, and baseline CD4 cell count. We used the estimated regression coefficients to impute an indicator of viral suppression where it was missing 100 times, estimated the overall proportion suppressed in each imputed dataset, and combined results across imputations using Rubin's rules.[18] In a separate analysis, we applied inverse probability of missingness weights[19] to upweight patients with a measured viral load to represent all patients meeting the eligibility criteria (including those with and without a measured viral load), based on the same measured covariates described above. Both approaches required the assumption that viral loads were missing at random conditional on the variables measured in the national database, but they relied on complementary modeling assumptions (imputation requires correct specification of the model for suppression, while weighting requires correct specification of the model for missingness).

Fourth, we used viral load measurements from the supplemental sample to augment the measurements in the national database.[20,21] To implement this approach, we upweighted participants in the supplemental sample to represent all eligible patients who were missing viral load data in the national database. Specifically, patients who had a viral load measurement in the national database received a weight of 1 (i.e., represented only themselves). Participants in the supplemental sample received a weight of y/n_0, where y was the total number of patients in the routine database without a recent viral load measurement and n_0 was the number in the supplemental sample. Finally, patients without a viral load measurement in the routine data or supplemental sample were given a weight of 0. The proportion of people on ART with a suppressed viral load was estimated in this weighted dataset. This approach is illustrated in Figure 1. We expect this approach to provide an unbiased estimate of the "third 95" if patients in the supplemental sample are representative of all patients missing viral load data in the national database. Theoretical justification for this approach is provided in eAppendix 1; http://links.lww.com/EDE/B941.

RESULTS

Of the 49,281 individuals registered in the national database, 21,517 entered care between 1 June 2013 and 25 January 2017, and 9,703 patients had started ART and were retained in care on 25 January 2017. These 9,703 patients constitute the target population for estimating viral suppression in this study. Only 26% (n = 2,548) of these patients had a viral load measurement recorded in the national database in the 6 months before 25 January 2017.

Nearly half (49%) of people on ART and retained in care on the target date were male, the majority were from the Dominican Republic (85%), and 50% were over age 40 (Table). Because the target population was limited to those retained in care on 25 January 2017, it was composed of more people who entered care in later years compared with earlier years. For example, 32% of eligible people entered care in 2016 compared with 14% in 2013. Baseline CD4 cell count at entry into care was missing for about two-thirds of people in the target population (66%); of available CD4 cell count



$$\text{Weights } W = \begin{cases} 1 & \text{For patients with recent viral load data} \\ \frac{\text{Number missing viral load data}}{\text{Number in validation sample}} & \text{For patients in the validation sample} \\ 0 & \text{For patients missing recent viral load data and not in the validation sample} \end{cases}$$

FIGURE 1. Illustration of the proposed two-stage approach to estimate the proportion of patients on treatment with a suppressed viral load in the Dominican Republic, 2017.

measurements at entry into care, the median was 262 cells/mm^3 (interquartile range, 119–413).

Distributions of measured patient characteristics were similar between people with and without a recent viral load measurement, although those missing recent viral load data were somewhat more likely to have been of Haitian nationality than those with recent viral load data. Measured characteristics of participants in the rapid assessment were similar to characteristics of people in the target population.

In the complete-case analysis, which was limited to the 2,548 patients with routinely collected viral load measurements in the national database, the estimated proportion suppressed was 60% (95% confidence interval [CI] = 58, 62), well below the 95–95–95 goal (Figure 2). Due to extensive missing data, the logical bounds on this estimate were wide; if all patients missing viral load information had an unsuppressed viral load, the proportion suppressed would be 16%, and if all patients missing viral load information had a suppressed viral load, the proportion suppressed would be 89%. After applying standard statistical approaches to account for missing data, results were similar to the complete-case analysis: the estimated proportion suppressed was 61% (95% CI = 59, 63) after using multiple imputation to account for missing data, and 61% (95% CI = 59, 63) after using inverse probability of missingness weights to account for missing data.

Among participants in the supplemental sample alone, the estimated proportion suppressed was 83% (95% CI = 79, 86), much higher than among those with routinely measured viral load information in the national database, but imprecise. Combining data from the supplemental sample with the routinely collected data using the two-stage approach, the estimated proportion suppressed was 77% (95% CI = 74, 80). Results were similar when limiting to patients enrolled in care at the nine selected sites, when extending the viral load assessment window from 6 to 12 months, and when statistical adjustment methods used a broader set of covariates (eAppendix 2; http://links.lww.com/EDE/B941).

DISCUSSION

We have illustrated a rapid sampling-based approach to estimating the "third 95" in clinical settings with incomplete viral load ascertainment. In our example from the Dominican Republic, the proportion of people with a suppressed viral load in the supplemental sample was much higher than the proportion with a suppressed viral load in routinely collected clinic records in a national database, suggesting that patients with suspected treatment failure may have been preferentially referred for the viral load testing recorded in clinic records. Incorporating viral load measurements from the supplemental sample resulted in an absolute increase of 17% in the estimated proportion suppressed, likely altering knowledge about the impact of the HIV care and treatment program.

Large population-based studies remain the gold standard for assessing population-level viral suppression at a specific point in time. However, such surveys are resource intensive and time consuming, and therefore are challenging to conduct with

TABLE. Characteristics of 9,703 People Who Entered HIV Care and Treatment in the Dominican Republic Between 1 June 2013 and 25 January 2017 and Were on ART and Retained in Care on 25 January 2017, and 452 Participants in a Supplemental Sample Missing Recent Routinely Collected Viral Load Data

Characteristic	Target Population, n = 9,703, n (%)	With Recent VL Data, n = 2,548, n (%)	Without Recent VL Data, n = 7,155, n (%)	Supplemental Sample, n = 452, n (%)
Male	4,771 (49)	1,309 (51)	3,462 (48)	211 (47)
Age at entry into care				
18–29	1,812 (19)	386 (15)	1,426 (10)	79 (18)
30–39	3,046 (31)	807 (32)	2,239 (31)	146 (32)
40–49	2,586 (27)	717 (28)	1,869 (26)	117 (26)
50 or over	2,259 (23)	638 (25)	1,621 (23)	110 (24)
Year of entry into care				
2013	1,448 (15)	353 (14)	1,095 (15)	64 (14)
2014	2,366 (24)	573 (23)	1,793 (25)	100 (22)
2015	2,773 (29)	732 (29)	2,041 (29)	146 (32)
2016	3,089 (32)	885 (35)	2,204 (31)	142 (31)
2017	27 (0)	5 (0)	22 (0)	0 (0)
Nationality				
Dominican Republic	8,279 (85)	2,262 (89)	6,017 (84)	399 (88)
Haiti	1,348 (14)	260 (10)	1,088 (15)	50 (11)
Other	76 (1)	26 (1)	50 (1)	3 (1)
Baseline CD4 cell count				
Missing	6,398 (66)	1,627 (64)	4,471 (67)	321 (71)
Under 200	1,263 (13)	351 (14)	912 (13)	49 (11)
200–349	907 (9)	232 (9)	675 (9)	39 (9)
350–500	608 (6)	190 (8)	418 (6)	23 (5)
Over 500	527 (5)	148 (6)	379 (5)	20 (4)

VL indicates viral load.

the frequency needed to monitor trends in real time. Clinical data, on the other hand, are generated over the course of routine healthcare provision, and, accordingly, are more likely to be "up to date." Moreover, individual HIV care and treatment programs may wish to assess trends in viral suppression among patients on ART in that specific program, rather than among people with HIV in a national or subnational administrative region.

As we have illustrated, a major limitation to using routinely collected clinical data to monitor trends in viral suppression in resource-limited settings is that biomarkers like viral load are often missing from clinical records. Such missingness is unlikely to occur at random; viral load may be more likely to be measured for patients appearing to fail treatment or seeking care in specific types of facilities. Moreover, as HIV care and treatment programs scale up viral load monitoring,[22–24] these patterns of missingness may change, such that a greater proportion of patients receive viral load measurements regularly, regardless of symptoms. If using only clinical records with available viral load measurements to estimate program-wide viral suppression, these shifts in who is receiving viral load measurements could induce an apparent trend in viral suppression over time, even if the proportion suppressed in the target population remains constant.

Here, we proposed a two-stage approach in which viral loads were proactively measured for a subset of patients with missing viral load data in routine clinic records. This approach does require resources beyond those needed in routine clinical data collection to sample patients for the rapid assessment and conduct additional viral load assays. However, such two-stage and "double sampling"[15] approaches have proven useful in many settings related to HIV. For example, researchers have used such designs to better estimate mortality and retention in care using clinical data in which outcomes like mortality, loss to care, and transfers to other clinics after loss to follow-up are missing.[16,25–29] These methods typically involve tracing a subset of those lost to follow-up from a specific clinical database to assess their outcomes and have been successfully applied to account for biases due to missing information after loss to follow-up in many resource-limited settings. Implementing the approach proposed here to account for missing viral load data is even more straightforward, requiring only sampling from patients who remain retained in HIV care and obtaining additional viral load measurements. With the proliferation of point of care viral load assays, rapid assessments like that described here could be conducted without reliance on a central laboratory, which may help facilitate use of this approach in routine practice settings.

In the example from the Dominican Republic, applying standard analytic approaches to account for missing data

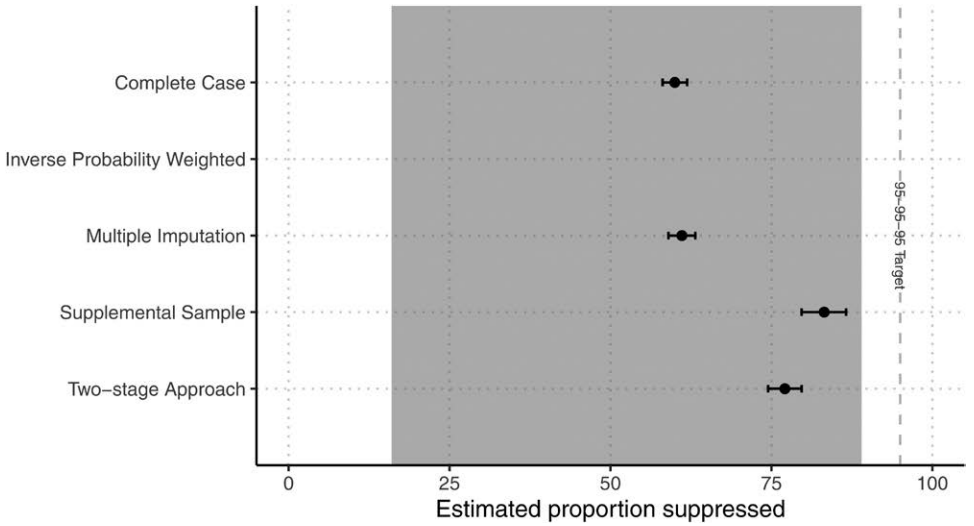


FIGURE 2. Estimates of the proportion of people with a suppressed viral load and comparison to UNAIDS 95–95–95 target among 9,703 people on ART and retained in HIV care in the Dominican Republic on 25 January 2017. Gray shading represents the plausible range of estimates based on logical bounds on the proportion suppressed given viral load data captured through routine clinical care.

(i.e., multiple imputation and inverse probability weighting) yielded an estimated proportion with a suppressed viral load very similar to the estimate from the complete-case analysis. These methods standardize the distribution of measured variables in the subset of the data with nonmissing data to match the distribution of these variables in the target population. These methods are expected to yield unbiased results if participants with and without data on the incompletely measured variable(s) are exchangeable conditional on measured covariates.[13] However, databases used to track clinical indicators in resource-limited settings may not include the complete set of covariates needed for such conditional exchangeability. Specifically, some covariates that are strongly associated with viral load, such as prior diagnoses and other markers of health status, are likely to be omitted from these databases. In the example from the Dominican Republic, we saw that the distribution of measured covariates was very similar between those with and without viral load measurements. Therefore, it was unsurprising that applying multiple imputation and inverse probability weighting to account for missing viral load data yielded estimates very close to the estimate from the complete-case analysis. Other approaches, such as linking multiple types of administrative datasets (e.g., laboratory data, clinical data, and vital records), may offer additional covariate data when available.

In settings without rich covariate data, analytic approaches alone may not account for bias due to missing viral load data. In these settings, extra information on the distribution of viral loads among those with viral load measurements missing from the routinely collected data is necessary to estimate the proportion suppressed in the full target population. Here, we obtained this information by conducting a rapid assessment in which we measured viral loads on a sample of people who were missing routinely collected viral load data in the national database. However, under a sufficient set of assumptions, one could also obtain this information from external sources (e.g., an assessment conducted among a similar, but external, population) or using expert knowledge. Alternatively, one could assess impact of missing data on estimates of viral suppression by estimating the proportion suppressed under various assumptions about the probability of suppression among those with missing viral load data using quantitative bias analysis.[30]

Using our approach, people with viral load measurements in the routinely collected data represent only themselves, while participants in the supplemental sample are upweighted to represent everyone in the target population (i.e., those on treatment and retained in care) without a routinely collected viral load measurement. This crude weighting approach is expected to yield valid estimates if participants in the supplemental sample are a random subset of those missing routinely collected viral load data. However, due to logistical constraints, we constructed the supplemental sample from the set of consecutive patients arriving at selected clinics during the data collection window as they were recruited by the rapid assessment. Because the target population was limited to patients retained in care, limiting the supplemental sample to those arriving at the clinic was unlikely to induce bias. However, if we oversampled people who came to the clinic more often than the average person retained in care, our estimates could have been biased in either direction. To assess if this was likely, we compared the average time between visits for those in the supplemental study (1.08 months) to the time between visits for those missing routinely collected viral load

measurements not in the supplemental sample (1.16 months). Because the visit intervals appeared similar, we do not expect that our approach to recruiting participants for the rapid assessment (and thus the supplemental sample) oversampled those with greater access to care.

We also relied on the assumption that individuals in the supplemental sample recruited from among those missing routine viral loads at the nine selected sites could stand in for people missing routine viral loads at all sites across the country. While we selected sites with wide geographic coverage and a range of characteristics (e.g., small and large; public and NGO), estimated viral suppression may not be valid at the national level if viral suppression among those missing viral load data at the nine selected sites differed from other sites. Finally, our study had a high response rate, but the validity of assessments like the one proposed here may be compromised if sampled individuals decline participation. If the probability of participation varies by covariates, the weights proposed here may be adapted to account for these covariates (eAppendix 1; http://links.lww.com/EDE/B941).

In this example, access to a national database with routinely collected data meant that we could determine which patients were in care at the time of the study with minimal concerns about "silent transfers" between health facilities within the country.[31] However, in settings without such national coverage, the supplemental sample may need to trace participants who appear to be out of care in the routinely collected records to accurately estimate viral suppression in the target population.

With the uptake of electronic records systems in many settings, routinely collected clinical data provide a wealth of information about both individual patients and population-level metrics in settings related to HIV and beyond. However, routinely collected clinical data are often plagued by the issues described here for viral load. Importantly, the presence or absence of data on a particular variable may be affected by the value of the variable, and missingness itself may affect health outcomes by limiting the information available for decisions about care.[32] The approach outlined here could be usefully applied in settings beyond HIV viral load analyses to leverage clinical data for learning about population health parameters while accounting for such bias due to missing data.

Improving knowledge in settings with imperfect data requires tradeoffs. Here, we have proposed an approach that requires limited additional data collection to account for bias due to missing viral load data. This approach provides insight into the distribution of viral suppression among people with HIV who are retained in care and missing routinely collected viral load data. In resource-constrained settings with suboptimal viral load coverage or gradually scaling up routine viral load monitoring,[33] implementing a sampling approach to estimate viral suppression could allow more accurate and efficient monitoring of HIV treatment program effectiveness.

REFERENCES

1. Mugavero MJ, Napravnik S, Cole SR, et al; Centers for AIDS Research Network of Integrated Clinical Systems (CNICS) Cohort Study. Viremia copy-years predicts mortality among treatment-naive HIV-infected patients initiating antiretroviral therapy. *Clin Infect Dis*. 2011;53:927–935.
2. Lesko CR, Edwards JK, Moore RD, Lau B. A longitudinal, HIV care continuum: 10-year restricted mean time in each care continuum stage after enrollment in care, by history of IDU. *AIDS*. 2016;30:2227–2234.
3. Gouskova NA, Cole SR, Eron JJ, Fine JP. Viral suppression in HIV studies: combining times to suppression and rebound. *Biometrics*. 2014;70:441–448.
4. Edwards JK, Cole SR, Adimora A, Fine J, Martin J, Eron J. Illustration of a measure to combine viral suppression and viral rebound in studies of HIV therapy. *J Acquir Immune Defic Syndr*. 2015;68:241–244.
5. Medland NA, McMahon JH, Chow EP, Elliott JH, Hoy JF, Fairley CK. The HIV care cascade: a systematic review of data sources, methodology and comparability. *J Int AIDS Soc*. 2015;18:20634.
6. Mugavero MJ, Amico KR, Westfall AO, et al. Early retention in HIV care and viral load suppression: implications for a test and treat approach to HIV prevention. *J Acquir Immune Defic Syndr*. 2012;59:86–93.
7. Quinn TC, Wawer MJ, Sewankambo N, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. *N Eng J Med*. 2000;342:921–929.
8. Joint United Nations Programme on HIV/AIDS and Joint United Nations Programme on HIV/AIDS. 90-90-90: An Ambitious Treatment Target to Help End the AIDS Epidemic. 2014. Available at: https://www.unaids.org/sites/default/files/media_asset/90-90-90_en.pdf. Accessed June 3, 2022.
9. Gaolathe T, Wirth KE, Holme MP, et al. Botswana' s progress toward achieving the 2020 UNAIDS 90-90-90 antiretroviral therapy and virological suppression goals: a population-based survey. *Lancet HIV*. 2016;3:221–230.
10. Marinda E, Simbayi L, Zuma K, et al. Towards achieving the 90-90-90 HIV targets: results from the south African 2017 national HIV survey. *BMC Public Health*. 2020;20:1375.
11. Brown K, Williams DB, Kinchen S, et al. Status of HIV epidemic control among adolescent girls and young women aged 15-24 years - seven African countries, 2015-2017. *MMWR Morb Mortal Wkly Rep*. 2018;67:29–32.
12. Lyatuu GW, Mwashemele SZ, Urrio R, et al. Long-term virological outcomes in women who started option B+ care during pregnancy for prevention of mother-to-child transmission of HIV in Dar es Salaam, Tanzania: a cohort study. *Lancet HIV*. 2021;8:e256–e265.
13. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. 2nd ed. Wiley-Interscience; 2002.
14. White JE. A two stage design for the study of the relationship between a rare exposure and a rare disease. *Am J Epidemiol*. 1982;115:119–128.
15. Frangakis CE, Rubin DB. Addressing an idiosyncrasy in estimating survival curves using double sampling in the presence of self-selected right censoring. *Biometrics*. 2001;57:333–342.
16. Geng EH, Emenyonu N, Bwana MB, Glidden DV, Martin JN. Sampling-based approach to determining outcomes of patients lost to follow-up in antiretroviral therapy scale-up programs in Africa. *JAMA*. 2008;300:506–507.
17. Eisinger RW, Dieffenbach CW, Fauci AS. HIV viral load and transmissibility of HIV infection: undetectable equals untransmittable. *JAMA*. 2019;321:451–452.
18. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. Wiley; 1987.
19. Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. *Stat Methods Med Res*. 2013;22:278–295.
20. Edwards JK, Zadrozny S, Hileman S, Seamans M, Goldberg S, Donastorg Y. *Improving Estimates of the HIV Care and Treatment Cascade among Key Populations in the Dominican Republic*. FHI360; 2017.
21. Zadrozny S, Weir S, Edwards JK, Herce M. *Applying New Methods to Estimate Viral Suppression: The "Last 90."* MEASURE Evaluation, University of North Carolina; 2018.
22. Alemnji G, Onyebujoh P, Nkengasong JN. Improving laboratory efficiencies to scale-up HIV viral load testing. *Curr Opin HIV AIDS*. 2017;12:165–170.
23. Roberts T, Cohn J, Bonner K, Hargreaves S. Scale-up of routine viral load testing in resource-poor settings: current and future implementation challenges. *Clin Infect Dis*. 2016;62:1043–1048.

24. Lecher S, Williams J, Fonjungo PN, et al. Progress with scale-up of HIV viral load monitoring - seven sub-Saharan African countries, January 2015-June 2016. *MMWR Morb Mortal Wkly Rep*. 2016;65:1332–1335.
25. Anderegg N, Johnson LF, Zaniewski E, et al; IeDEA, MeSH consortia. All-cause mortality in HIV-positive adults starting combination antiretroviral therapy: correcting for loss to follow-up. *AIDS*. 2017;31(suppl 1):S31–S40.
26. Brinkhof MW, Pujades-Rodriguez M, Egger M. Mortality of patients lost to follow-up in antiretroviral treatment programmes in resource-limited settings: systematic review and meta-analysis. *PLoS One*. 2009;4:e5790.
27. Egger M, Spycher BD, Sidle J, et al; IeDEA East Africa, West Africa and Southern Africa. Correcting mortality for loss to follow-up: a nomogram applied to antiretroviral treatment programmes in sub-Saharan Africa. *PLoS Med*. 2011;8:e1000390.
28. Chammartin F, Zürcher K, Keiser O, et al. Outcomes of patients lost to follow-up in African antiretroviral therapy programs: individual patient data meta-analysis. *Clin Infect Dis*. 2018;67:1643–1652.
29. Vanobberghen F, Weisser M, Kasuga B, et al. Mortality rate in a cohort of people living with HIV in rural Tanzania after accounting for unseen deaths among those lost to follow-up. *Am J Epidemiol*. 2021;190:251–264.
30. Lash TL, Fox MP, Fink AK. *Applying Quantitative Bias Analysis to Epidemiologic Data*. Springer Science & Business Media; 2011.
31. Edwards JK, Lesko CR, Herce ME, et al. Gone but not lost: implications for estimating HIV care outcomes when loss to clinic is not loss to care. *Epidemiology*. 2020;31:570–577.
32. Caniglia EC, Sabin C, Robins JM, et al; Center for AIDS Research Network of Integrated Clinical Systems and the HIV-CAUSAL Collaboration. When to monitor CD4 cell count and HIV RNA to reduce mortality and AIDS-defining illness in virologically suppressed HIV-positive persons on antiretroviral therapy in high-income countries: a prospective observational study. *J Acquir Immune Defic Syndr*. 2016;72:214–221.
33. Roberts T, Cohn J, Bonner K, Hargreaves S. Scale-up of routine viral load testing in resource-poor settings: current and future implementation challenges: table 1. *Clin Infect Dis*. 2016;62:1043–1048.