



Practice of Epidemiology

Laboratory Measures as Proxies for Primary Care Encounters: Implications for Quantifying Clinical Retention Among HIV-Infected Adults in North America

Peter F. Rebeiro*, Keri N. Althoff, Bryan Lau, John Gill, Alison G. Abraham, Michael A. Horberg, Mari M. Kitahata, Baligh R. Yehia, Hasina Samji, John T. Brooks, Kate Buchacz, Sonia Napravnik, Michael J. Silverberg, Anita Rachlis, Kelly A. Gebo, Timothy R. Sterling, Richard D. Moore, and Stephen J. Gange for the North American AIDS Cohort Collaboration on Research and Design

* Correspondence to Dr. Peter F. Rebeiro, Division of Infectious Diseases, Department of Medicine, Vanderbilt University School of Medicine, 1161 21st Avenue South, A-2200 Medical Center North, Nashville, TN 37232 (e-mail: p.rebeiro@vanderbilt.edu).

Initially submitted August 31, 2014; accepted for publication June 29, 2015.

Because of limitations in the availability of data on primary care encounters, patient retention in human immunodeficiency virus (HIV) care is often estimated using laboratory measurement dates as proxies for clinical encounters, leading to possible outcome misclassification. This study included 83,041 HIV-infected adults from 14 clinical cohorts in the North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD) who had ≥ 1 HIV primary care encounters during 2000–2010, contributing 468,816 person-years of follow-up. Encounter-based retention (R_{EB}) was defined as ≥ 2 encounters in a calendar year, ≥ 90 days apart. Laboratory-based retention (R_{LB}) was defined similarly, using the dates of CD4-positive cell counts or HIV-1 RNA measurements. Percentage of agreement and the κ statistic were used to characterize agreement between R_{LB} and R_{EB} . Logistic regression with generalized estimating equations and stabilized inverse-probability-of-selection weights was used to elucidate temporal trends and the discriminatory power of R_{LB} as a predictor of R_{EB} , accounting for age, sex, race/ethnicity, primary HIV risk factor, and cohort site as potential confounders. Both R_{EB} and R_{LB} increased from 2000 to 2010 (from 67% to 78% and from 65% to 77%, respectively), though R_{EB} was higher than R_{LB} throughout ($P < 0.01$). R_{LB} agreed well with R_{EB} (80%–86% agreement; $\kappa = 0.55$ – 0.62 , $P < 0.01$) and had a strong, imperfect ability to discriminate between persons retained and not retained in care by R_{EB} (C statistic: $C = 0.81$, $P < 0.05$). As a proxy for R_{EB} , R_{LB} had a sensitivity and specificity of 84% and 77%, respectively, with misclassification error of 18%.

clinical encounters; clinical retention; HIV; laboratory measurements; measurement error; misclassification; proxies

Abbreviations: AIDS, acquired immunodeficiency syndrome; CI, confidence interval; HIV, human immunodeficiency virus; HIV-1, HIV type 1; IDU, injection drug user; IPW, inverse-probability-of-selection weights; NA-ACCORD, North American AIDS Cohort Collaboration on Research and Design; ROC, receiver operating characteristic; ROR, ratio of odds ratios.

Access to health care and its utilization for clinical monitoring in both healthy populations (for preventive care) and unhealthy populations (for treatment) are vital components of improving health outcomes. The reliable and ongoing use of clinical services by patients as part of disease management, or retention in care, is one such indicator of health-care utilization. Epidemiologic tools have long been recognized as being both available and useful for evaluation of evidence and formulation of policies for health services (1). With health-care reform (e.g., the Affordable Care Act) removing barriers

to linkage to care in the United States, an understanding of how people seek health care and are retained in care is likely to become increasingly valuable (2, 3). These questions speak to identifying disparities in our health-care system and may have a large impact on patient outcomes.

The study of contemporaneous human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) epidemiology in the United States has increasingly recognized health-care retention as a key component in improving health outcomes in individuals and decreasing transmission of HIV.

It is a core indicator of quality care for HIV-infected persons and a central feature of what has been termed the “HIV care continuum” (4–7). In 2010, the National HIV/AIDS Strategy advocated increasing the proportion of HIV-infected patients in continuous care (8); and in 2012, the Institute of Medicine outlined measures for assessing both the National HIV/AIDS Strategy and benchmarks outlined in the Affordable Care Act (9). More recently, the Department of Health and Human Services adopted process indicators for HIV care (10), and the President of the United States issued an Executive Order directing the Office of National AIDS Policy to coordinate a federal response to improve engagement across the HIV care continuum (11, 12).

Amid heightened national emphasis on improving clinical retention among persons with HIV infection, correctly quantifying retention at the national, state, and local levels is essential to accurately assess progress toward established benchmarks. However, measures of retention in HIV care have varied depending on data availability and the populations being studied, with various demographic, clinical, and environmental characteristics having been shown to influence care patterns over time (13–16). Laboratory data have been widely used as surrogate measures when data on primary care encounters are unavailable, as is often the case in large surveillance studies using public health laboratory data or cross-sectional population-based surveys that sample across defined geographic areas (e.g., county or state surveys). Such measures have been shown to be good indicators of initial or overall access to clinical HIV care (17), but there are no longitudinal data regarding concordance between measures of retention defined by HIV primary care encounters and those using surrogate laboratory measures (6, 9, 18, 19).

With widespread accessibility and use of laboratory measures adapted as indicators of ongoing clinical care, there is a need to examine agreement across these measures, which are based on varying data sources, so that policy-makers, researchers, and other consumers of epidemiologic information can better understand and interpolate the clinical care experience indicated by surrogate measures. Therefore, in this study we quantified concordance between laboratory measures (used as surrogates for direct measurement of clinical encounters) and clinical encounters in a North American consortium of AIDS cohort studies.

METHODS

Population and study design

The North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD) is a large demographically, clinically, and geographically diverse collaboration which has been collecting data from multi- and single-site interval and clinical cohort studies in the United States and Canada since 2006 (20). According to the Institute of Medicine, the NA-ACCORD has demonstrated a constituent patient population that comprises a large proportion of and is demographically similar to persons living with HIV/AIDS in the United States, and it has been recognized as one of the data systems adequate to assess and quantify quality-of-care measures (including retention in clinical care) that will serve

as benchmarks for progress under the National HIV/AIDS Strategy and the Affordable Care Act (9). Further details on the NA-ACCORD collaboration have been published previously (20).

Clinical and demographic data from 25 cohort studies contributing information to NA-ACCORD (including laboratory test values and their collection dates, medical diagnoses and dates, antiretroviral drugs used and prescription dates, clinical encounter data, insurance status, and the first 3 digits of the zip code of residence) are transmitted annually to a centrally administered Data Management Core, where all contributed data sets are harmonized. Data undergo quality control for completeness and accuracy, including measures to reduce the probability that an individual is concurrently participating in more than 1 clinical cohort study. The activities of both the NA-ACCORD centrally and each participating cohort study have been reviewed and approved by their respective local institutional review boards.

Among clinical cohorts, only patients who made 2 or more clinic visits within 12 months were enrolled in the NA-ACCORD, limiting the NA-ACCORD clinical population to patients established as being “in care.” For this analysis, we examined adult participants who had 1 or more HIV primary care encounters between January 2000 and December 2010. Interval cohorts, with visits conducted outside of clinical care activities (21), were excluded to allow an exclusive focus on patterns of patient clinical care. The 14 included clinical cohorts were comprised of patients from all 50 US states; Washington, DC; Puerto Rico; and 9 of the 13 Canadian provinces and territories.

Retention measures, factors associated with retention, and follow-up

This study compared 2 strategies for quantifying clinical retention. The first was encounter-based clinical retention (R_{EB}), defined as the Institute of Medicine-based indicator: ≥ 2 HIV primary care encounters within each calendar year, ≥ 90 days apart (9). The second was laboratory-based retention (R_{LB}), defined in the same fashion as R_{EB} but using dates of CD4-positive T-lymphocyte (CD4+ cell) count or HIV type 1 (HIV-1) RNA measurement rather than primary care encounters as markers of care. Clinical systems submitting data for these analyses, such as Kaiser Permanente Northern California, the Veterans Aging Cohort Study, and others, extracted data on HIV primary care encounters directly from electronic health records or from the billing system, both of which create an encounter record only if the patient visit actually occurred; that is, these data reflect a visit only if the patient appeared for care, not just scheduled an appointment. Because all cohort studies represented primary care settings, not urgent care, these were adjudged to be true primary care encounters. Beyond HIV primary care encounters, those coded as subspecialty care, emergency department, dental, mental health, nursing, nutrition, orientation, pharmacy, social work/case management, substance abuse, or other/unknown outpatient encounters were excluded. Due to strong financial incentives, our clinical cohorts would require highly accurate assessments of patient encounters for reimbursement purposes, in either billing or electronic health records systems. Therefore,

Table 1. Percentage of Person-Time Retained in Care for HIV-Positive Patients as Defined by Both Primary Care Encounters and Laboratory Measures, According to Demographic, Clinical, and Geographic Characteristics, NA-ACCORD, 2000–2010^a

Factor	No. of Patients	% ^b	% of Person-Time Retained ^c		Agreement Between R_{LB} and R_{EB}	
			R_{EB}	R_{LB}	κ	%
Total	83,041	100	71	67	0.59	80
Age, years						
<40	35,713	43	62	63	0.64	83
40–49	29,542	36	70	67	0.61	83
50–59	13,863	17	77	68	0.53	81
≥60	3,923	5	84	70	0.39	78
Sex						
Male	67,951	82	71	66	0.57	80
Female	15,090	18	69	69	0.57	80
Race/ethnicity						
Non-Hispanic white	33,207	40	72	69	0.56	82
Non-Hispanic black	34,175	41	69	62	0.60	82
Hispanic	9,510	11	72	73	0.64	86
Other/unknown	6,149	7	70	67	0.59	82
Primary HIV risk factor						
Men who have sex with men	30,589	37	69	74	0.63	85
Injection drug user	14,329	17	69	58	0.58	80
Heterosexual contact	18,908	23	68	71	0.65	85
Other/unknown	19,215	23	76	59	0.49	77
CD4-positive T-lymphocyte count, cells/mm ³ ^d						
<200	19,322	29	70	69	0.64	85
200–349	14,718	22	73	74	0.63	86
350–499	12,940	20	74	75	0.61	85
≥500	19,061	29	75	76	0.59	85
HIV-1 RNA level, copies/mL ^d						
<200	16,972	27	81	83	0.52	86
≥200	45,205	73	69	68	0.63	84
Receipt of ART for ≥6 months/year ^e						
No ART	46,449	56	50	42	0.57	79
ART	36,592	44	83	82	0.46	84
Country of care						
United States	79,236	95	70	66	0.58	82
Canada	3,805	5	74	75	0.66	87

Abbreviations: ART, antiretroviral therapy; HIV, human immunodeficiency virus; HIV-1, HIV type 1; NA-ACCORD, North American AIDS Cohort Collaboration on Research and Design.

^a Encounter-based retention (R_{EB}) was defined as ≥2 primary care encounters in a calendar year, ≥90 days apart. Laboratory-based retention (R_{LB}) was defined similarly, using the dates of CD4-positive cell counts or HIV-1 RNA measurements.

^b Percentages may not sum to 100 because of rounding.

^c The percentage of person-time retained in care during the study according to the encounter-based measure (R_{EB}) (i.e., the percentage of years spent “in care” between cohort entry and the final encounter) differed from the percentage of person-time retained according to the laboratory-based measure (R_{LB}) within every stratum (χ^2 test: $P < 0.01$).

^d Data on CD4-positive cell count ($n = 66,041$) and HIV-1 RNA level ($n = 62,177$) at cohort entry were not available for all participants.

^e Receipt of ART (≥3 agents from ≥2 classes or a triple nucleoside/nucleotide reverse transcriptase inhibitor regimen containing abacavir or tenofovir) during the year of cohort entry.

any error in capturing the dates of patient HIV primary care encounters—very basic functions in these systems—should have been negligible and more likely random than systematic in nature, though no additional validation was performed. Therefore, no bias should have been introduced by their use as described in this analysis. Inpatient data were also excluded.

Data from individuals were used to create 1 observation per year between the year of the person's entry into the cohort and the year of his/her final encounter or laboratory measurement prior to the end of 2010. This single annual observation carried a summary of the individual's retention status, depending on whether his/her data for that year met criteria for R_{EB} or R_{LB} (see Web Appendix 1 and Web Figure 1, available at <http://aje.oxfordjournals.org/>). The initial year of care in the cohort was excluded if the patient entered in the final quarter of a calendar year (and was thus ineligible to be "retained" in his/her year of entry into care). The year of death was excluded for patients who died before the end of the study (due to individuals not being uniformly at risk for the outcome in the year of death, dependent on the timing of their death during the calendar year). Follow-up time ranged between a minimum of 1 year and a maximum of 11 years, and individuals could contribute multiple observations over the study period.

Participant age (categorized as <40 years, 40–49 years, 50–59 years, or ≥ 60 years), sex, self-characterized race/ethnicity (categorized as white, black, Hispanic, or other/unknown), self-characterized primary risk factor for HIV acquisition (categorized as men who have sex with men, injection drug user (IDU), heterosexual contact, or other/unknown), receipt of antiretroviral therapy for ≥ 6 months in a year (≥ 3 antiretroviral agents from ≥ 2 classes, or a triple nucleoside/nucleotide reverse transcriptase inhibitor regimen containing abacavir or tenofovir), CD4+ cell count (categorized as <200, 200–349, 350–499, or ≥ 500 cells/mm³), HIV-1 RNA level (categorized as <200 or ≥ 200 copies/mL), and site of clinical care were included in descriptive analyses as factors by which the agreement between R_{EB} and R_{LB} may have differed. Receipt of antiretroviral therapy, CD4+ cell count, and HIV-1 RNA level were excluded from regression analyses because of their potential to induce bias as time-dependent confounders of the relationship between R_{LB} and R_{EB} . CD4+ cell counts were the first values measured in the calendar year of cohort entry, and HIV-1 RNA levels were the last values measured; the choice of relative timing for these measures was related to widely used indicators of clinical care (9). All factors were time-varying except sex, race/ethnicity, and primary HIV risk factor.

Statistical models and methods

Percentage of agreement, discordance (percentage of negative agreement), sensitivity, specificity, and the kappa statistic (κ) were all used to quantify agreement between R_{EB} and R_{LB} across and within demographic and clinical characteristics (22). Differences between R_{EB} and R_{LB} within strata of baseline characteristics were detected by χ^2 test. Weighted logistic regression using a generalized estimating equation was used to assess temporal trends and construct receiver operating characteristic (ROC) curves with their respective C statistics

to assess the discriminatory power of R_{LB} as a predictor of R_{EB} (23, 24). A Toeplitz correlation structure based on the mean values of the unstructured covariance for repeated outcomes within individuals was used in the generalized estimating equation regression (25, 26).

Construction of stabilized inverse probability weights

Inverse-probability-of-selection weights (IPW) are an increasingly popular strategy for addressing the problem of confounding. By reweighting (or standardizing) populations with respect to the exposure conditional on measured potential confounders, we can eliminate the exposure-confounder link when estimating the effect or association of the exposure with the outcome (27). In this analysis, we used these techniques to determine the marginal relationship between R_{EB} and R_{LB} while accounting for the potential confounders baseline age, sex, race/ethnicity, primary HIV risk factor, and cohort site in the regression model (28). These were time-fixed confounders, so we used single weights across all time points for an individual. The weights were also stabilized and truncated at the 5th and 95th percentiles to improve balance (29).

The weights themselves were constructed using regression to estimate the probability of exposure, conditioning on the appropriate confounding factors. Because the exposure of interest was R_{LB} , a dichotomous variable, a logistic regression model was used to create the IPW. Additional details on model diagnostics used in the selection and construction of the Toeplitz correlation structure for the regression with a generalized estimating equation (Web Appendix 2, Web Figures 2 and 3, Web Table 1) and the distribution of untruncated

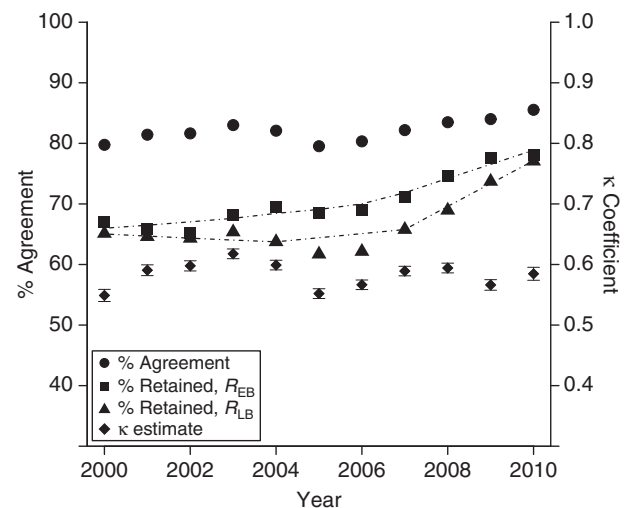


Figure 1. Temporal trends in encounter- and laboratory-based measures of retention in care (R_{EB} and R_{LB} , respectively) and agreement between measures (% agreement and κ coefficient) in the North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD), 2000–2010. Dashed-and-dotted lines across R_{EB} and R_{LB} percentages are locally weight scatterplot smoothing (LOESS) curves for trend. κ estimates (diamonds) are plotted with 95% confidence intervals (bars).

and truncated IPW accounting for site alone and for all available confounders (Web Appendix 3, Web Figure 4) are given in the Web appendices. All analyses were performed using R, version 3.0.3 (R Foundation for Statistical Computing, Vienna, Austria) (www.r-project.org).

RESULTS

Among 83,041 adults included in the study population, the median amount of time contributed was 4 years (interquartile range, 2–6 years), with a total contribution of 468,816 person-years. There were significant differences in R_{EB} versus R_{LB} across all patient demographic and clinical factors (Table 1). The person-time retained ranged between 60% and 80% for both R_{EB} and R_{LB} for most characteristics but was lower among persons with a primary HIV risk factor of IDU and those not receiving antiretroviral therapy for ≥ 6 months per year. Despite these differences, according to accepted epidemiologic measures, there was good agreement between R_{EB} and R_{LB} within all demographic and clinical strata (77%–87% agreement; $\kappa = 0.39$ – 0.66 , all P 's < 0.01) (Table 1).

There were increases in retention throughout the study period according to both definitions, though retention was higher when it was measured by R_{EB} throughout ($P < 0.01$ for trend; Figure 1). For R_{EB} , the proportion of patients retained increased from 67% (20,591/30,741) in 2000 to 78% (26,701/34,205) in 2010. Similarly, for R_{LB} , 65% (20,020/30,741) and 77% (26,357/34,205) of patients were retained in 2000 and 2010, respectively (Figure 1). There was fair agreement between R_{EB} and R_{LB} over time (80%–86% agreement; $\kappa = 0.55$ – 0.62 , $P < 0.01$) (Figure 1).

Regression models utilizing a generalized estimating equation, with and without stabilized IPW to account for potential variations in practice across clinical sites (R_{EB} ranged from 62% to 85% across sites and R_{LB} from 51% to 79%) and differences across age, sex, race/ethnicity, and HIV risk categories, were used to construct ROC curves. Using the area under the ROC curve for the weighted model, R_{LB} had a strong but imperfect ability to discriminate between persons retained in care and those not retained by R_{EB} ($C = 0.805$, $P < 0.05$) (Figure 2A). Using the area under the ROC curve for the robust adjusted model (incorporating IPW to adjust for site, age, sex, race/ethnicity, and primary HIV risk factor), the discrimination of R_{LB} was improved slightly over that of the unweighted model ($C = 0.868$, $P < 0.05$) (Figure 2B). Using predicted probabilities of R_{EB} from the regression models plotted as thresholds along the ROC curves, the sensitivities and specificities of R_{LB} as a proxy for R_{EB} were maximized at a retention probability of 30% in the marginal weighted model and at retention probabilities of 60% to 80% in the adjusted weighted model, which was in the range of observed retention levels over the study period (Figure 2A and 2B). Using weighted model estimates, the sensitivity and specificity of R_{LB} as a proxy for R_{EB} in the pseudopopulation were 84% and 77%, respectively, which resulted in a discordance of 18% (Table 2). Results were similar when using predicted probabilities from the weighted regression model and shorter intervals during the study period (2000–2003, 2004–2007, and 2008–2010), though discordance according to R_{LB} was slightly higher in earlier periods (18% and 19%, respectively) than in the most recent period (16%). Exclusion of the largest cohort in the study sample, which had reduced barriers to

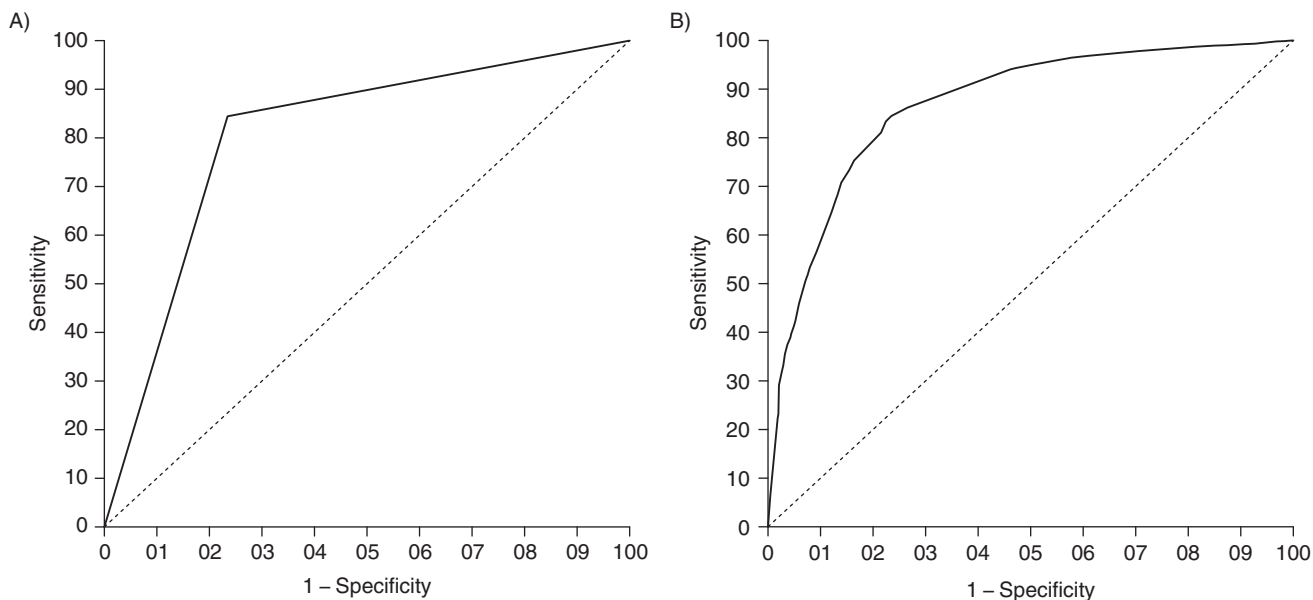


Figure 2. Receiver operating characteristic (ROC) curves quantifying the discrimination of an encounter-based measure of retention of human immunodeficiency virus (HIV)-positive patients in care (R_{EB}) by a laboratory-based measure of retention in care (R_{LB}), North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD), 2000–2010. The curves were derived from regression models with generalized estimating equations and inverse-probability-of-selection weights (IPW). A) ROC curve with IPW based on site, age, sex, race/ethnicity, and primary HIV risk factor. B) ROC curve derived from a model with robust adjustment for clinic site with IPW. Predicted probabilities of retention by R_{EB} are thresholds on the ROC curves.

Table 2. Agreement Between Encounter-Based Measures (R_{EB}) and Laboratory-Based Measures (R_{LB}) of Retention in Care for HIV-Positive Patients and Discrimination of R_{EB} by R_{LB} , NA-ACCORD, 2000–2010^a

Method of Parameter Estimation	No. of Observations	C Statistic	% Discordant	Sensitivity, %	Specificity, %	κ
Weighted data using IPW	468,816		17	86	79	0.63
Regression model with IPW	468,816	0.87	18	84	77	0.58

Abbreviations: HIV, human immunodeficiency virus; IPW, inverse-probability-of-selection weights; NA-ACCORD, North American AIDS Cohort Collaboration on Research and Design.

^a Estimated from a model using stabilized IPW constructed using clinic site, age, sex, race/ethnicity, and primary HIV risk factor.

accessing care, resulted in R_{LB} 's being closer to R_{EB} throughout the study period, with only a slight elevation in the percentage of agreement and κ values (percent agreement $\leq 87\%$ and $\kappa \leq 0.67$; data not shown). In addition, the overall misclassification error (discordance) dropped slightly from 18% to 16%.

Application of the same approaches within subpopulations of special concern (e.g., younger, minority, female, and IDU patients) revealed similar patterns of improving retention over time. However, in weighted regression analysis with factor-by- R_{LB} joint-effect interactions, also including main effects, agreement between R_{EB} and R_{LB} was lower among persons over age 50 years versus persons aged ≤ 50 years (ratio of odds ratios (ROR) = 0.71, 95% confidence interval (CI): 0.67, 0.74) and IDU patients versus non-IDU patients (ROR = 0.91, 95% CI: 0.86, 0.96) and was higher among females versus males (ROR = 1.12, 95% CI: 1.06, 1.18) and black patients versus nonblack patients (ROR = 1.05, 95% CI: 1.01, 1.10) (Figure 3).

DISCUSSION

The accurate measurement of retention in care has implications for the epidemiology and management of HIV infection

in the United States and Canada, as well as the design, implementation, and evaluation of prevention and treatment strategies for HIV and other chronic diseases (8). It is also clear that measuring clinical retention accurately per policy guidelines and refining targets for improvement are consistent with the recent emphasis on a “test and treat strategy” to rapidly diagnose HIV-infected persons and engage them in continuous care, which has been theorized to “bend the incidence curve” of HIV in the United States and Canada downward (30, 31). Therefore, in this study we sought to leverage longitudinal clinical data available within a large, demographically and geographically diverse collaborative consortium of cohort studies to elucidate the critical and underexplored issue of misclassification when using laboratory proxies for clinical encounters in describing the epidemiology of clinical retention.

Even in a clinically engaged population who have successfully linked to health care, using laboratory measures as proxies for clinical encounters does not provide a complete picture of engagement in care as recommended by recently adopted indicators (9, 10, 32). The potential for misclassification of an encounter-based retention outcome based on laboratory-based retention measures is small but nontrivial, ranging from 15% discordance to 20% discordance. Moreover, we

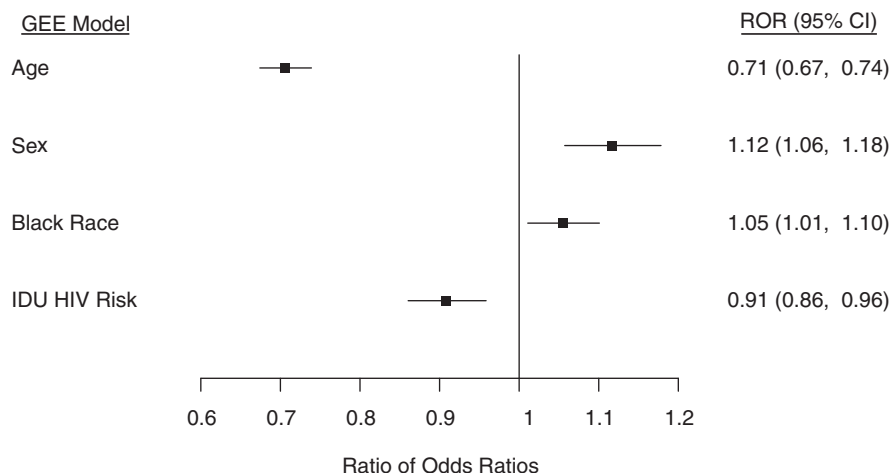


Figure 3. Ratios of odds ratios (RORs) and 95% confidence intervals (CIs) for the probability of retaining human immunodeficiency virus (HIV)-positive patients in care as assessed by an encounter-based measure (R_{EB}), conditional on retention assessed by a laboratory-based measure (R_{LB}), among subpopulations of concern (joint-effects interactions), North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD), 2000–2010. Reference groups: for age, ≤ 50 years; for sex, male; for black race, nonblack race; for injection drug use (IDU) as the primary HIV risk factor, a non-IDU primary HIV risk factor. GEE, generalized estimating equations.

identified significant (although modest in magnitude) interactions between the level of agreement in these measures and membership in subpopulations of special concern in the HIV epidemic; this should alert epidemiologists to the possibility of differential misclassification.

There are many reasons R_{EB} may not agree perfectly with R_{LB} . For example, in one participating cohort study, initial laboratory panels were ordered at enrollment by medical care management personnel, prior to the patient's first clinical encounter. In other instances, clinics facilitated laboratory monitoring at regular intervals under standing orders from clinicians, without the requirement of a direct clinical encounter. Alternatively, patients may access laboratory testing at a facility that does not require clinical orders (though this may be relatively rare). In any of these cases, patients may meet the definition of retention by R_{LB} even if not by R_{EB} . The opposite may hold for patients receiving clinical care without laboratory monitoring at regular intervals.

Though the frequency of laboratory monitoring may be of great concern to a clinician or epidemiologist depending on the needs of the clinic population served, it is not equivalent to the frequency of clinic attendance, even at a less granular level (e.g., an annually assessed retention indicator), and may not serve the same policy or monitoring purposes despite its utility otherwise. This is denoted in the language of the Institute of Medicine report, which refers to laboratory measures as "proxies" for encounter-based measures (9, 17). The use of CD4+ cell counts and HIV-1 RNA measurements to denote retention may also be of concern, since monitoring frequency guidelines or state/local reporting practices for HIV-related laboratory measures may change in different populations (19, 33). In particular, as guidelines continue to decrease the recommended frequency of laboratory monitoring among stable patients on antiretroviral therapy, the degree to which laboratory-based measures will underestimate the population of retained patients will increase. Thus, it is not obvious that one methodology is superior to the other; quality measures of timely laboratory measurement and of timely encounters may both have merit.

Additionally, we are not asserting that assessing retention using laboratory proxies is a worthless endeavor (the perfect ought not be the enemy of the good). Neither are encounter-based retention measures the "gold standard" for measuring the clinical experience of patients (the engagement of patients in their own care and resulting improved HIV disease progression may well be a mixture of interactions with clinicians and receipt of ancillary services); there could also be limitations in complete capture of all encounter data when formulating R_{EB} measures. However, when policies or strategies are structured around indicators defined using a particular measure, monitoring the effects of those policies with fidelity to their design may be limited by reliance on imperfect proxies. The use of laboratory measures as surrogates for primary care encounters by public health agencies is, after all, an exercise in maximizing the utility of readily available data, and while statistically different they are within reasonable range of each other. Researchers and policy-minded professionals interested in assessing population-level changes in health-care service indicators using these data must be aware of the constraints their estimates face.

As it stands, not all CD4+ cell counts or HIV-1 RNA measurements are uniformly reported across jurisdictions to the

National HIV Surveillance System; currently, mandatory reporting of all values is the legal standard in 42 of the 50 US states and Washington, DC, although, as of 2014, only 18 jurisdictions (representing 52% of adults diagnosed with HIV in the United States) were considered to have "complete" reporting sufficient to include in annual surveillance estimates (34). The Centers for Disease Control and Prevention has implemented the Medical Monitoring Project, which collects both clinical encounter and HIV laboratory information to supplement the National HIV Surveillance System, but the Medical Monitoring Project is limited by its cross-sectional design (35). When examining retention at the population level, then, there may be few options for researchers and policy analysts other than employing proxy measures, though these data may be incomplete indicators of R_{EB} . Again, a discordance of 15%–20% between R_{EB} and R_{LB} is plausible based on the results from our large cohort. Though evaluation of possible differences in agreement between R_{EB} and R_{LB} by state of patient residence was outside the scope of the current analysis, it is also possible that incomplete data due to differences in reporting practices could be ameliorated and retention itself improved by the implementation of secure health information exchanges or other emerging data-sharing solutions (36, 37). Such exchanges, resulting in unified health records, could potentially greatly improve surveillance activities for health-care service outcomes, even beyond HIV.

That being said, it is important to recognize that the results of this analysis do not show a large divergence between R_{EB} and R_{LB} . For monitoring purposes, it may be the case that use of one versus the other may result in different judgments about whether a particular benchmark for clinical retention has been met, but the size of the misclassification error comparing one with the other is low enough that the distinction may not be as important above a particularly high benchmark. For example, the National HIV/AIDS Strategy establishes a goal to "increase the proportion of Ryan White HIV/AIDS Program clients who are in continuous care . . . from 73% to 80%" (8, p. 21). While previous years showed greater discordance, results from 2010, with 77% of patients successfully retained according to R_{LB} , may not be meaningfully different for either policy or clinical purposes from a population that is 78% retained by R_{EB} (8). Whatever the difference in different populations, though, it remains important to quantify it and to speak clearly and openly about the fact that these two measures are not completely interchangeable and, in the future, may or may not become even more divergent.

There were limitations in this analysis due to characteristics of the study population and the statistical traits of some measures of agreement. Because of the high prevalence of retention in this population, the κ statistic in particular may have returned artificially low estimates of agreement; this may also occur when the prevalence of the outcome is very low (22). However, multiple measures of agreement and predictive discrimination were used to achieve a more complete picture of the relationship between R_{EB} and R_{LB} over time in this population. Another potential limitation of this analysis is that results obtained using the NA-ACCORD patient population may not fully represent the continuum of care experienced by all patients, particularly if they access clinical care outside the network of sites in NA-ACCORD, though

NA-ACCORD participant sites comprise a wide array of clinical settings across the United States and Canada—county hospitals, private practices, health maintenance organizations, Veterans Affairs facilities, etc. Further, our results appeared sensitive to the exclusion of our largest cohort, though levels of retention and intermeasure agreement were quite similar to those in the overall cohort, using weighted analyses. This may suggest the influence of individual cohort characteristics on agreement between measures, though this is expected if clinical sites have very different models of care. Because all data in NA-ACCORD are anonymized before being harmonized at the central Data Management Core, there is no way of tracking the movement of individuals between cohort sites (though the large geographic dispersion of sites makes it likely that individuals are not simultaneously accessing care at disparate locations); there may be misclassification error in the outcome if patients leave care at a member site and access care through a local public health department, private physician's office, unaffiliated local hospital, or other venue, because they may appear to be experiencing suboptimal retention during that period even though they are not. However, this would be a shortcoming in any setting where there is not a comprehensive, nationally linked medical records or claims database. Despite these potential limitations, NA-ACCORD has a very large sample size that is demographically representative of persons living with HIV/AIDS in the United States, represents persons living in geographically diverse regions of the United States and Canada, and has been formally endorsed as an ideal data source with which to assess progress under the National HIV/AIDS Strategy, which is an exercise related to just the sorts of issues addressed above (9, 38).

In summary, this analysis showed that clinical retention by both R_{EB} and R_{LB} improved over time and that R_{EB} and R_{LB} were strongly correlated. However, our findings suggest that R_{LB} was an imperfect surrogate for R_{EB} , and we were able to quantify the level of imperfection. The discordance between encounter retention status and widely used laboratory proxies provides further motivation for the development of novel health information-sharing strategies and structures to facilitate the most accurate assessment of HIV clinical retention possible.

ACKNOWLEDGMENTS

Author affiliations: Division of Infectious Diseases, Department of Medicine, School of Medicine, Vanderbilt University, Nashville, Tennessee (Peter F. Rebeiro, Timothy R. Sterling); Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland (Keri N. Althoff, Bryan Lau, Alison G. Abraham, Kelly A. Gebo, Richard D. Moore, Stephen J. Gange); Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada (John Gill); Mid-Atlantic Permanente Research Institute, Kaiser Permanente Mid-Atlantic States, Rockville, Maryland (Michael A. Horberg); Division of Allergy and Infectious Diseases, Department of Medicine, School of Medicine, University of Washington, Seattle, Washington (Mari M.

Kitahata); Division of Infectious Diseases, Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania (Baligh R. Yehia); Epidemiology and Population Health Program, British Columbia Centre for Excellence in HIV/AIDS, Vancouver, British Columbia, Canada (Hasina Samji); Division of HIV/AIDS Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia (John T. Brooks, Kate Buchacz); Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (Sonia Napravnik); Kaiser Permanente Division of Research, Kaiser Permanente Northern California, Oakland, California (Michael J. Silverberg); and Division of Infectious Diseases, Department of Medicine, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada (Anita Rachlis).

All authors contributed equally to this work.

This work was supported by grants U01-AI069918, U01-AA013566, U24-AA020794, U01-AA020790, U01-AI31834, U01-AI34989, U01-AI34993, U01-AI34994, U01-AI35004, U01-AI35039, U01-AI35040, U01-AI35041, U01-AI35042, U01-AI35043, U01-AI37613, U01-AI37984, U01-AI38855, U01-AI38858, U01-AI42590, U01-AI68634, U01-AI68636, U01-AI69432, U01-AI69434, U01-HD32632, U10-EY08057, U10-EY08052, U10-EY08067, U11-RR024131, U11-TR000083, U54-MD007587, F31-DA035713, G12-MD007583, K01-AI071754, K01-AI093197, K23-MH097647, K23 EY013707, K24-DA00432, MO1-RR-00052, N02-CP55504, P30-AI027763, P30-AI094189, P30-AI27757, P30-AI27767, P30-AI50410, P30-AI54999, P30-AI036219, P30-MH62246, R01-CA165937, R01-AA16893, R01-DA11602, R01-DA04334, R01-DA12568, R24-AI067039, R56-AI102622, Z01-CP010214, and Z01-CP010176 from the National Institutes of Health, US Department of Health and Human Services; contract CDC200-2006-18797 from the Centers for Disease Control and Prevention, US Department of Health and Human Services; contract 90047713 from the Agency for Healthcare Research and Quality, US Department of Health and Human Services; contract 90051652 from the Health Resources and Services Administration, US Department of Health and Human Services; grants TGF-96118, HCP-97105, CBR-86906, and CBR-94036 from the Canadian Institutes of Health Research; the Ontario Ministry of Health and Long Term Care; and the Government of Alberta, Canada.

We are grateful to all physicians, investigators, and staff involved in NA-ACCORD.

This work was presented in preliminary form at the 18th International Workshop on HIV Observational Databases, Sitges, Spain, March 27–29, 2014.

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the US Centers for Disease Control and Prevention.

Conflict of interest: none declared.

REFERENCES

1. Armenian HK, Shapiro S. *Epidemiology and Health Services*. New York, NY: Oxford University Press; 1998.

2. Galea S. An argument for a consequentialist epidemiology. *Am J Epidemiol*. 2013;178(8):1185–1191.
3. Cates W Jr. Invited commentary: consequential(ist) epidemiology: let's seize the day. *Am J Epidemiol*. 2013;178(8):1192–1194.
4. Gardner EM, McLees MP, Steiner JF, et al. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. *Clin Infect Dis*. 2011;52(6):793–800.
5. Giordano TP, Gifford AL, White AC Jr, et al. Retention in care: a challenge to survival with HIV infection. *Clin Infect Dis*. 2007;44(11):1493–1499.
6. Mugavero MJ, Davila JA, Nevin CR, et al. From access to engagement: measuring retention in outpatient HIV clinical care. *AIDS Patient Care STDS*. 2010;24(10):607–613.
7. Horberg MA, Aberg JA, Cheever LW, et al. Development of national and multiagency HIV care quality measures. *Clin Infect Dis*. 2010;51(6):732–738.
8. White House Office of National AIDS Policy. *National HIV/AIDS Strategy for the United States*. Washington, DC: White House Office of National AIDS Policy; 2010. <http://purl.access.gpo.gov/GPO/LPS124282>. Accessed June 24, 2015.
9. Ford MA, Spicer CM. *Monitoring HIV Care in the United States: Indicators and Data Systems*. Washington, DC: National Academies Press; 2012.
10. Valdiserri RO, Forsyth AD, Yakovchenko V, et al. Measuring what matters: development of standard HIV core indicators across the U.S. Department of Health and Human Services. *Public Health Rep*. 2013;128(5):354–359.
11. White House Office of National AIDS Policy. *Improving Outcomes: Accelerating Progress Along the HIV Care Continuum*. Washington, DC: White House Office of National AIDS Policy; 2013. http://www.whitehouse.gov/sites/default/files/onap_nhas_improving_outcomes_dec_2013.pdf. Accessed June 24, 2015.
12. National Archives and Records Administration. Executive Order 13649—accelerating improvements in HIV prevention and care in the United States through the HIV Care Continuum Initiative. *Fed Regist*. 2013;78(138):43057–43059.
13. Fleishman JA, Yehia BR, Moore RD, et al. Establishment, retention, and loss to follow-up in outpatient HIV care. *J Acquir Immune Defic Syndr*. 2012;60(3):249–259.
14. Giordano TP. Retention in HIV care: what the clinician needs to know. *Top Antivir Med*. 2011;19(1):12–16.
15. Howe CJ, Cole SR, Napravnik S, et al. Enrollment, retention, and visit attendance in the University of North Carolina Center for AIDS Research HIV clinical cohort, 2001–2007. *AIDS Res Hum Retroviruses*. 2010;26(8):875–881.
16. Yehia BR, French B, Fleishman JA, et al. Retention in care is more strongly associated with viral suppression in HIV-infected patients with lower versus higher CD4 counts. *J Acquir Immune Defic Syndr*. 2014;65(3):333–339.
17. Perkins D, Meyerson BE, Klinkenberg D, et al. Assessing HIV care and unmet need: eight data bases and a bit of perseverance. *AIDS Care*. 2008;20(3):318–326.
18. Centers for Disease Control and Prevention. Vital signs: HIV prevention through care and treatment—United States. *MMWR Morb Mortal Wkly Rep*. 2011;60(47):1618–1623.
19. Hall HI, Gray KM, Tang T, et al. Retention in care of adults and adolescents living with HIV in 13 U.S. areas. *J Acquir Immune Defic Syndr*. 2012;60(1):77–82.
20. Gange SJ, Kitahata MM, Saag MS, et al. Cohort profile: the North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD). *Int J Epidemiol*. 2007;36(2):294–301.
21. Lau B, Gange SJ, Moore RD. Interval and clinical cohort studies: epidemiological issues. *AIDS Res Hum Retroviruses*. 2007;23(6):769–776.
22. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;37(5):360–363.
23. Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73(1):13–22.
24. Harrell FE. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York, NY: Springer Publishing Company; 2001.
25. Yan J, Fine J. Estimating equations for association structures. *Stat Med*. 2004;23(6):859–874.
26. Højsgaard S, Halekoh U, Yan J. The R package geepack for generalized estimating equations. *J Stat Softw*. 2006;15(2):1–11.
27. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.
28. van der Wal WM, Geskus RB. Ipw: an R package for inverse probability weighting. *J Stat Softw*. 2011;43(i13):1–23.
29. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168(6):656–664.
30. Nachege JB, Uthman OA, del Rio C, et al. Addressing the Achilles' heel in the HIV care continuum for the success of a test-and-treat strategy to achieve an AIDS-free generation. *Clin Infect Dis*. 2014;59(suppl 1):S21–S27.
31. Marks G, Gardner LI, Craw J, et al. Entry and retention in medical care among HIV-diagnosed persons: a meta-analysis. *AIDS*. 2010;24(17):2665–2678.
32. 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(1):86–93.
33. Dean BB, Hart RL, Buchacz K, et al. HIV laboratory monitoring reliably identifies persons engaged in care. *J Acquir Immune Defic Syndr*. 2015;68(2):133–139.
34. Gale HB, Gitterman SR, Hoffman HJ, et al. Is frequent CD4+ T-lymphocyte count monitoring necessary for persons with counts ≥ 300 cells/ μ L and HIV-1 suppression? *Clin Infect Dis*. 2013;56(9):1340–1343.
35. Centers for Disease Control and Prevention. Monitoring selected national HIV prevention and care objectives by using HIV surveillance data, United States and 6 US dependent areas, 2010. *HIV Surveill Suppl Rep*. 2012;17(3):1–27.
36. Magnus M, Herwehe J, Gruber D, et al. Improved HIV-related outcomes associated with implementation of a novel public health information exchange. *Int J Med Inform*. 2012;81(10):e30–e38.
37. Virga PH, Jin B, Thomas J, et al. Electronic health information technology as a tool for improving quality of care and health outcomes for HIV/AIDS patients. *Int J Med Inform*. 2012;81(10):e39–e45.
38. Althoff KN, Buchacz K, Hall HI, et al. U.S. trends in antiretroviral therapy use, HIV RNA plasma viral loads, and CD4 T-lymphocyte cell counts among HIV-infected persons, 2000 to 2008. *Ann Intern Med*. 2012;157(5):325–335.