



Practice of Epidemiology

Measurement of Current Substance Use in a Cohort of HIV-Infected Persons in Continuity HIV Care, 2007–2015

Catherine R. Lesko*, Alexander P. Keil, Richard D. Moore, Geetanjali Chander, Anthony T. Fojo, and Bryan Lau

* Correspondence to Dr. Catherine R. Lesko, Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205 (e-mail: clesko2@jhu.edu).

Initially submitted July 23, 2017; accepted for publication April 18, 2018.

Accurate, routine measurement of recent illicit substance use is challenging. The Johns Hopkins Human Immunodeficiency Virus Clinical Cohort (Baltimore, Maryland) collects 2 imperfect but routine measurements of recent substance use: medical record review and self-interview. We used Bayesian latent class modeling to estimate sensitivity and specificity of each measurement as well as prevalence of substance use among 2,064 patients engaged in care during 2007–2015. Sensitivity of medical record review was higher than sensitivity of self-interview for cocaine and heroin use; posterior estimates ranged from 44% to 76% for cocaine use and from 39% to 67% for heroin use, depending on model assumptions and priors. In contrast, sensitivity of self-interview was higher than sensitivity of medical record review for any alcohol use, hazardous alcohol use, and cigarette smoking. Posterior estimates of sensitivity of self-interview were generally above 80%, 85%, and 87% for each substance, respectively. Specificity was high for all measurements. From one model, we estimated prevalence of substance use in the cohort to be 12.5% for cocaine, 9.3% for heroin, 48.5% for alcohol, 21.4% for hazardous alcohol, and 55.4% for cigarettes. Prevalence estimates from other models were generally comparable. Measurement error of substance use is nontrivial and should be accounted for in subsequent analyses.

measurement error; patient-reported outcome measures; prevalence; sensitivity; specificity; substance use

Abbreviations: CI, confidence interval; CrI, credible interval; DIC, deviance information criterion; HIV, human immunodeficiency virus; JHHCC, Johns Hopkins HIV Clinical Cohort; MRR, medical record review; SI, self-interview.

Illicit drug use, alcohol misuse, and cigarette smoking are associated with poor health outcomes among persons with human immunodeficiency virus (HIV) (1–12). Yet time-updated substance use data are not widely collected or reported in clinical care settings. Most studies of current drug use in persons with HIV are characterized by small convenience samples. As a result, associations between substance use and HIV outcomes and estimated substance-use prevalence may be biased (13, 14). The Johns Hopkins HIV Clinical Cohort (JHHCC) is one of few clinical cohorts to regularly assess substance use; assessment measures include medical record review (MRR) and self-interview (SI) (15, 16). As such, the JHHCC is uniquely positioned to describe substance use among patients receiving routine HIV care. However, underreporting of substance use by the physician in the medical record, or by the patient on a self-administered survey, may bias

estimates of prevalence and of the effects of substance use if measurement error is not addressed (17–19).

Our objective was to characterize the performance of routinely implemented MRR and SI for assessing recent use of cocaine, heroin, any alcohol, hazardous alcohol, and cigarettes. We used Bayesian latent class analysis to estimate the sensitivity and specificity of these 2 assessments as well as the prevalence of substance use in the subsample of patients included in this analysis.

METHODS

Study sample

The JHHCC enrolls all HIV-infected adults (≥ 18 years of age) receiving continuity care at Johns Hopkins HIV clinic who

consent to share their data (20). Patient characteristics, including sex, race, age, and HIV acquisition risk factors are abstracted from medical record documentation. Collection of data on patients in the cohort and their analysis were approved by the Johns Hopkins Hospital institutional review board.

The study sample included 2,064 persons engaged in care between January 1, 2007, and August 31, 2015, who had both a completed MRR and SI. We randomly selected 1 person-month per person during this follow-up period for analysis. Models fitted on all person-months available for each patient did not yield meaningfully different results (data not shown).

Ascertainment of current substance use

Medical record review. Cocaine, heroin, alcohol (hazardous/social/none), and cigarette use in the past 6 months are abstracted from patients' medical records every 6 months. The majority of indications of substance use in the medical record are physician progress notes; however, drug and alcohol screening test results and diagnoses of addiction or alcohol use disorder are also abstracted.

Self-interview. A subset of patients completes a computer-assisted SI approximately every 6 months, coinciding with a clinic visit. During the study period, SI was not offered to everyone due to limited resources, and no particular type of patient was targeted for participation. Because clinic visits do not occur exactly every 6 months, in contrast to the MRR, SI does not cover the entire person-time in care for patients who have completed at least 1 SI. Approximately two-thirds of patients have completed at least 1 SI. On the SI, patients reported cocaine, heroin, and cigarette use (yes/no) in the past 6 months and the quantity and frequency of alcohol use (21). We defined hazardous alcohol use on the SI as: 1) ≥ 7 drinks per week for women or ≥ 14 drinks per week for men; or 2) any report of binge drinking (defined as ≥ 6 drinks on 1 occasion from 2007 to September 2013 (22), subsequently reduced to ≥ 4 drinks on 1 occasion for women or ≥ 5 drinks on 1 occasion for men) (23). SI results are not shared with patients' providers and are not part of the medical record.

Because the 6-month periods covered by the MRR and SI do not overlap by design, we split patients' follow-up time into 1-month periods for analysis to capture partial overlap. We assigned values for substance use to the 6 person-months prior to and including the date of the MRR or SI.

Statistical analysis

In the absence of a gold-standard assessment of substance use, there is not enough information to estimate sensitivity or specificity of MRR or SI or prevalence of substance use without making further assumptions (24). In order to evaluate the robustness of our results to different assumptions, we provide results under various assumptions regarding the accuracy of measurement or prevalence of substance use.

With these issues in mind, we combined 2 approaches to estimating all parameters of interest (sensitivity and specificity of MRR and SI and prevalence of substance use): 1) Split the population into groups that vary according to prevalence of substance use but across which sensitivity and specificity of MRR and SI can reasonably be assumed to be identical (25); and 2) apply a Bayesian framework using knowledge about sensitivity or specificity

derived from external information or expert opinion (26). Both approaches allow for estimation of sensitivity and specificity of measurements and prevalence of substance use in the absence of a gold standard. Assuming SI and MRR sensitivities and specificities are conditionally independent given true substance use, all parameters of interest are identifiable under the first approach. Combining the first approach with the Bayesian framework allowed us to characterize uncertainty about parameter estimates (24). Furthermore, incorporating a Bayesian framework allowed us to relax some of the assumptions of the first approach: that sensitivities of the 2 measurements were identical across groups (model 2 below) and that sensitivities of the 2 measurements were independent conditional on the presence of substance use (model 3 below) (27, 28).

We stratified the sample into 16 subgroups defined by the cross-classification of sex (male vs. female), race (black vs. non-black), age (<45 years vs. ≥ 45 years; 45 years was close to median age at enrollment into the cohort), and HIV acquisition risk factor (injection drug use vs. anything else) assumed to have different prevalence of substance use. Analytically, our approach involves assuming there is a latent indicator of true substance use, so the method has previously been referred to as Bayesian latent class analysis (29–32).

For each model described below, for each substance, we estimated the sensitivity and specificity of MRR, the sensitivity and specificity of SI, and the prevalence of substance use. We fitted the models described below using the statistical software JAGS and the R (R Foundation for Statistical Computing, Vienna, Austria) package "rjags" (33, 34) and the statistical software Stan and the R package "rstan" (35, 36). Code is available in Web Appendix 1 (available at <https://academic.oup.com/aje>). For JAGS models, we generated 4 parallel, independent Markov chain Monte Carlo samples from different starting values, each with 50,000 iterations after a 10,000 iteration burn-in. Stan models converged more quickly, so we generated 3 parallel, independent Markov chain Monte Carlo samples from different starting values, each with 5,000 iterations after a 5,000 iteration burn-in. We assessed convergence using visual inspection of trace plots and Rhat statistics. We report medians of the last 50,000 (5,000 for Stan models) samples and use the 2.5th and 97.5th percentiles from the distribution of samples for reporting the 95% posterior credible interval.

Models. To provide an anchor for estimates from subsequent models that increasingly relax assumptions about the specificity of the 2 measures of substance use and about the conditional independence of their sensitivities, we first estimated sensitivities of MRR and SI and prevalence of substance use under the implausibly strong assumption that specificity of both measures was perfect (i.e., assuming no false positives with 100% certainty). We call this model 0. This can be thought of as a special case of the Bayesian latent class analysis with a point-mass prior for specificity on 100%. Next, we assumed that sensitivities and specificities for both MRR and SI were identical across subgroups (model 1) and later relaxed the assumption of equal sensitivities of MRR and SI and allowed for different sensitivities for MRR and SI within each of the 16 subgroups defined by sex, race, age, and HIV acquisition risk factor (model 2). Because one could reasonably expect specificities of MRR and SI to be high (and this expectation was confirmed in model 1), we decided a priori that variations in specificity across subgroups would not be substantively meaningful, and we assumed constant specificity of both measures. We

compared the fit of model 2 with that of model 1 for each substance using the deviance information criterion (DIC) (33, 37) to determine whether allowing separate sensitivity parameters fit the data better than estimating a single sensitivity for the entire sample. If model 2 had had a lower DIC than model 1, we intended to collapse some of the 16 groups with similar sensitivity of MRR or SI and compare the DIC of the simplified model with that of the full model 2 iteratively, until we identified the model with the optimal number of unique sensitivity parameters. In practice, model 2 did not fit better than model 1 for any substance (sensitivity was not meaningfully different across subgroup).

To relax the assumption of conditionally independent sensitivities, we introduced a covariance term between the 2 measurements' sensitivities that was constrained to be positive (model 3). Allowing the covariance term to span the entire plausible range, given the sensitivities, did not meaningfully alter the results (data not shown). Frequentist approaches to this model would require at least 4 different measurements to estimate all parameters of interest uniquely (27, 28). In the absence of additional measurements, we needed to constrain one more parameter (in addition to constraints on specificities); we felt that the most prior knowledge was available on the plausible range of the prevalence of each substance use and incorporated weak priors on prevalence to ensure model 3 converged. Because, in most cases, results from model 3 were substantively different from model 1 (in contrast to model 2 versus model 1), and because different priors were required to estimate model 3 versus model 1, we did not compare models 3 and 1 using DIC. Rather, we report results from both models.

Priors. Prior distributions were elicited from physicians working in the HIV clinic by asking them to give their opinion about the plausible range (95% credible interval (CrI)) and most likely values for sensitivity and specificity of MRR and SI. Generally, physicians were unsure about the sensitivity, but they thought specificity for both measures would be very high—and higher for SI than for MRR. Thus, for most models, we assumed a beta(10, 90) prior on the specificity of MRR (median, 90.3% specificity, 95% CrI: 83.5, 95.0) and a beta(5, 95) prior on the specificity of SI (median, 95.3%, 95% CrI: 90.0, 98.3). Because we could not elicit informative priors for sensitivity, for most models we assumed uniform(0,1) priors on sensitivity of either measure. As described above, we needed to constrain one additional parameter in model 3, which included a term for covariance between the sensitivity of the MRR and SI for person-periods in which there was substance use; we chose to put a weak prior on prevalence of substance use. We explored the influence of the choice of priors on the posterior probability density functions by refitting nearly all models with different priors; a full list of models/priors fit is available in Table 1.

RESULTS

The majority of patients in our analytic sample were male (64%) and black (81%). Median age at start of analysis was 48 years (interquartile range, 41–54). A third (34%) reported history of injection drug use as their most probable route of HIV acquisition (Table 2).

Data on substance use from SI and MRR are shown in Table 3. We describe the data and results from models 1 and 3 for each substance below. Results from model 1 are illustrated in Figure 1.

In general, neither assessment was exceptionally sensitive for detecting illicit drug use. SI was most sensitive for detecting any alcohol, hazardous alcohol, and cigarette use. While our priors on specificity of MRR and SI were moderately informative, posteriors on specificity were not constrained by the priors; posterior estimates of specificity of MRR and SI for identifying recent cocaine and heroin use were higher than, and nonoverlapping with, the prior. Model 2 (which allowed for sensitivities of the 2 measures to vary by group) did not fit the data appreciably better for any substance (DIC of model 2 was higher than DIC of model 1); therefore, we do not discuss model 2 results, but abridged results appear in Table 4 for comparison, and full results appear in Web Table 1. Results from model 3 (which allowed for conditional dependence in sensitivities of MRR and SI) were dependent upon the prior for the prevalence of substance use; more details are available below and in Table 3.

Cocaine

Using a randomly selected person-month for members of the study sample, prevalence of recent cocaine use in the cohort as indicated on the MRR was 10.2%, while prevalence using the SI alone was 6.9%. MRR and SI both indicated use of cocaine use 4.0% of the time. If we assumed both MRR and SI had perfect specificity (model 0), estimated sensitivities of MRR and SI were 78.4% (95% confidence interval (CI): 73.5, 83.4) and 52.4% (46.4, 58.4), respectively. Based on model 1, estimated sensitivity was 68.8% (95% CrI: 55.7, 83.6) for MRR and 44.5% (95% CrI: 36.2, 53.6) for SI. When we relaxed the assumption of conditional independence of sensitivities, estimated sensitivity for each measure and estimated posterior prevalence of cocaine use varied based on the assumed prior probability of cocaine use. In general, posterior estimates of sensitivity of each measure were inversely related to the prior put on prevalence of cocaine use.

Heroin

Using a randomly selected person-month for members of the study sample, prevalence of heroin indicated on the MRR was 7.1%, while prevalence using the SI alone was 3.8%. MRR and SI both indicated heroin use 2.6% of the time. Under model 0, assuming perfect specificity, estimated sensitivities of MRR and SI were 84.9% (95% CI: 79.5, 90.2) and 45.9% (95% CI: 38.5, 53.4), respectively. Based on model 1, estimated sensitivities of MRR and SI were 66.8% (95% CrI: 53.7, 79.3) and 38.6% (95% CrI: 29.6, 48.5), respectively. When we relaxed the assumption of conditional independence of sensitivities, as was the case with recent cocaine use, estimated sensitivities for both measures were generally lower than in model 1 and inversely related to the prior put on prevalence of recent heroin use. Estimated specificities for heroin use were higher than estimated specificities for cocaine use for both measures and across all models.

Any alcohol

Prevalence of alcohol use for the cohort was 27.2% based on the MRR and 41.9% based on the SI. Both the MRR and SI indicated alcohol use for 17.4% of patients during a randomly selected person-month. Under model 0, estimated sensitivity of MRR and

Table 1. List of Prior Distributions and Models Explored in Investigation of the Performance of Medical Record Review and Self-Interview for Detecting Recent Substance Use Among Persons With Human Immunodeficiency Virus in Routine Medical Care, Johns Hopkins Human Immunodeficiency Virus Clinical Cohort, 2007–2015

Model	Sensitivity		Specificity		Prevalence
	MRR	SI	MRR	SI	
Model 1a ^a	uniform(0,1)	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	uniform(0,1)
Model 1b ^a	beta(2, 5) ^d	beta(2, 5) ^d	beta(10,0.5) ^e	beta(10,0.5) ^e	uniform(0,1)
Model 2 ^f	uniform(0,1)	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	uniform(0,1)
Model 3a (C,H,HA) ^g	uniform(0,1)	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	beta(1, 5) ^h
Model 3b ^g (C,H,HA)	uniform(0,1)	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	beta(1, 7) ⁱ
Model 3c ^g (C,H)	uniform(0,1)	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	beta(1, 3) ^j
Model 3c ^g (HA)	uniform(0,1)	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	beta(3, 9) ^k
Model 3d ^g (HA)	uniform(0,1)	beta(5, 95) ^c	beta(10, 90) ^b	beta(5, 95) ^c	beta(1, 3) ^l
Model 3a ^g (A)	uniform(0,1)	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	beta(2, 4) ^m
Model 3b ^g (A)	uniform(0,1)	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	beta(2, 6) ⁿ
Model 3c ^g (A)	uniform(0,1)	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	beta(2) ^o
Model 3d ^g (A)	beta(10, 90) ^b	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	beta(2) ^o
Model 3e ^g (A)	beta(5, 95) ^c	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	beta(2) ^o
Model 3a ^g (T)	uniform(0,1)	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	beta(3) ^p
Model 3b ^g (T)	uniform(0,1)	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	beta(2, 3) ^q
Model 3c ^g (T)	uniform(0,1)	uniform(0,1)	beta(10, 90) ^b	beta(5, 95) ^c	beta(2, 3) ^r

Abbreviations: A, any alcohol use; C, cocaine use; CrI, credible interval; H, heroin use; HA, hazardous alcohol use; MRR, medical record review; SI, self-interview; T, tobacco cigarette use.

^a Model 1 assumes conditional independence of sensitivities and specificities; assumes sensitivity and specificity of MRR and SI are constant across subgroups defined by patient demographics.

^b Median, 90.3% (95% CrI: 83.5, 95.0).

^c Median, 95.3% (95% CrI: 90.0, 98.2).

^d Median, 73.5% (95% CrI: 35.8, 95.7).

^e Median, 97.7% (95% CrI: 77.2, 100).

^f Model 2 assumes conditional independence of sensitivities and specificities; assumes specificities of MRR and SI are constant across subgroups defined by patient demographics but allows for subgroup-specific sensitivities of MRR and SI.

^g Model 3 assumes conditional independence of specificities but allows for conditional dependence of sensitivities; assumes sensitivities and specificities are constant across subgroups.

^h Median, 13.0% (95% CrI: 0.5, 52.2).

ⁱ Median, 9.4% (95% CrI: 0.4, 40.9).

^j Median, 20.6% (95% CrI: 0.8, 70.8).

^k Median, 23.6% (95% CrI: 6.0, 51.8).

^l Median, 20.6% (95% CrI: 0.8, 70.7).

^m Median, 31.4% (95% CrI: 5.3, 71.7).

ⁿ Median, 22.8% (95% CrI: 3.7, 58.0).

^o Median, 50.0% (95% CrI: 9.4, 93.3).

^p Median, 50.0% (95% CrI: 14.7, 85.4).

^q Median, 38.6% (95% CrI: 6.8, 80.6).

^r Median, 61.5% (95% CrI: 19.4, 93.3).

SI was 52.6% (95% CI: 49.6, 55.6) and 81.0% (95% CI: 78.6, 83.3). Under model 1, estimated sensitivity of MRR was lower, 41.7% (95% CrI: 38.3, 45.3), while estimated sensitivity of SI was slightly higher, 84.7% (95% CrI: 74.5, 96.6). Specificity of MRR was lower for detecting alcohol use than for all other substances. When we relaxed the assumption of conditional independence of sensitivities, posterior sensitivity of MRR was generally stable around 40%, regardless of the priors put on prevalence of alcohol use. Posterior sensitivity of SI was more dependent on the prior prevalence of alcohol use with estimates ranging from 75% to 97%.

Hazardous alcohol

Prevalence of hazardous alcohol use for the cohort was 11.5% based solely on the MRR and 17.5% based solely on the SI. Both the MRR and SI indicated hazardous alcohol use for 5.1% of patients during a randomly selected person-month. Under model 0, estimated sensitivities of MRR and SI were 48.2% (95% CI: 43.8, 52.6) and 73.3% (95% CI: 69.4, 77.2), respectively. Under model 1, estimated sensitivity of MRR was 29.2% (95% CrI: 24.3, 34.7). Posterior sensitivity of SI was 78.4% (95% CrI: 57.1, 98.0). Similar to any alcohol use, specificity of the MRR was suboptimal for

Table 2. Characteristics of 2,064 HIV-Persons Infected With Human Immunodeficiency Virus, With at Least 1 Medical Record Review and at Least 1 Self-Interview at the Time of the Randomly Selected Person-Month, Johns Hopkins Human Immunodeficiency Virus Clinical Cohort, 2007–2015

Patient Characteristic	No.	%
Male sex	1,318	64
Age ^a	48 (41–54)	
Black race	1,678	81
History of injection drug use	702	34
MSM	483	23
ART-exposed	1,903	92
Detectable viral load	481	23
CD4 cell count (cells/ μ L) ^a	441 (258–635)	

Abbreviations: ART, antiretroviral therapy (≥ 3 drugs initiated on the same day); HIV, human immunodeficiency virus; MSM, men who have sex with men.

^a Values are expressed as median (interquartile range).

hazardous alcohol use (93.1%, 95% CrI: 91.4, 95.2). Allowing the sensitivities of the 2 measurements to covary did not meaningfully increase the posterior sensitivity or specificity of the MRR. Estimates of the posterior sensitivity of SI for detecting recent hazardous alcohol use ranged from 65% to 93% depending on the prior put on prevalence of hazardous alcohol use.

Cigarettes

Prevalence of cigarette use indicated on the MRR was 46.3%, while prevalence using the SI alone was 49.6%. MRR and SI both indicated cigarette use 36.4% of the time. Under model 0, estimated sensitivities of MRR and SI were 77.8% (95% CI: 75.5, 80.1) and 83.3% (95% CI: 81.2, 85.4). Under model 1, estimated sensitivities of MRR and SI were 75.1% (95% CrI: 71.8, 78.9) and 87.2% (95% CrI: 82.9, 91.7), respectively. Relaxing the assumption of conditional independence of the sensitivities of MRR and SI did not meaningfully change posterior estimates

for sensitivity or specificity of either MRR or SI, nor did it meaningfully change estimates of prevalence of cigarette smoking.

DISCUSSION

In general, MRR or SI alone had poor sensitivity for detecting recent substance use in this cohort of persons with HIV engaged in continuity medical care. Assuming both measures were 100% specific (model 0; no false positives) tended to meaningfully overestimate sensitivity of the measurements as well as certainty of those estimates. All other models (with more diffuse priors on parameters) yielded arguably more realistic estimates of the performance of each measure for the detection of different substances; estimated sensitivity of the 2 measures were generally quite different from model 0 estimates (usually lower), and their precision is not overstated. When we allowed the sensitivity of MRR and SI to covary among patients assumed to have recently used heroin or cocaine, posterior estimates were susceptible to prior assumptions about the prevalence of recent use. This is unavoidable in the absence of more assessment measures (27, 28). If prevalence was assumed to be lower, estimated posterior sensitivities were higher, and vice versa. Ultimately, information from both measurements should be used to identify recent substance use as neither has sufficiently high sensitivity to be used alone.

Overall, although posterior sensitivity estimates varied, sometimes widely, we can conclude that: 1) MRR was more sensitive than SI for detecting recent cocaine and recent heroin use; 2) median posterior sensitivity of MRR for detecting recent cocaine use ranged from 44% to 76%, with the most likely estimate around 70%; 3) median posterior sensitivity of MRR for detecting recent heroin use ranged from 39% to 67%, with the most likely estimate around 60%; 4) SI was more sensitive than MRR for detecting recent alcohol use, recent hazardous alcohol use, or recent cigarette use; 5) sensitivities of SI for detecting recent alcohol, hazardous alcohol, and cigarette use are likely upwards of 80%, 85%, and 87%, respectively.

This is one of relatively few studies that have compared collection of sensitive information via self-report with data collection in the electronic health record. McGinnis et al. (38) found higher prevalence of alcohol use, and Kozak et al. (39) found higher prevalence of illicit substance use among persons in care for HIV

Table 3. Prevalence (Person-Months and Row Percent) of Current Substance Use as Reported on Self-Interview and Abstracted From Medical Record Review in 1 Randomly Selected Person-Month per 2,064 Persons in the Johns Hopkins Human Immunodeficiency Virus Clinical Cohort, 2007–2015

Substance	SI				No SI			
	MRR		No MRR		MRR		No MRR	
	Person-Months	%	Person-Months	%	Person-Months	%	Person-Months	%
Cocaine	83	4	58	3	128	6	1,795	87
Heroin	53	3	26	1	93	5	1,892	92 ^a
Alcohol	359	17	506	25	203	10	996	48
Hazardous alcohol	106	5	256	12	132	6	1,570	76 ^a
Cigarettes	751	36	273	13	205	10	835	40 ^a

Abbreviations: MRR, medical record review; SI, self-interview.

^a Row percentages do not sum to 100 due to rounding.

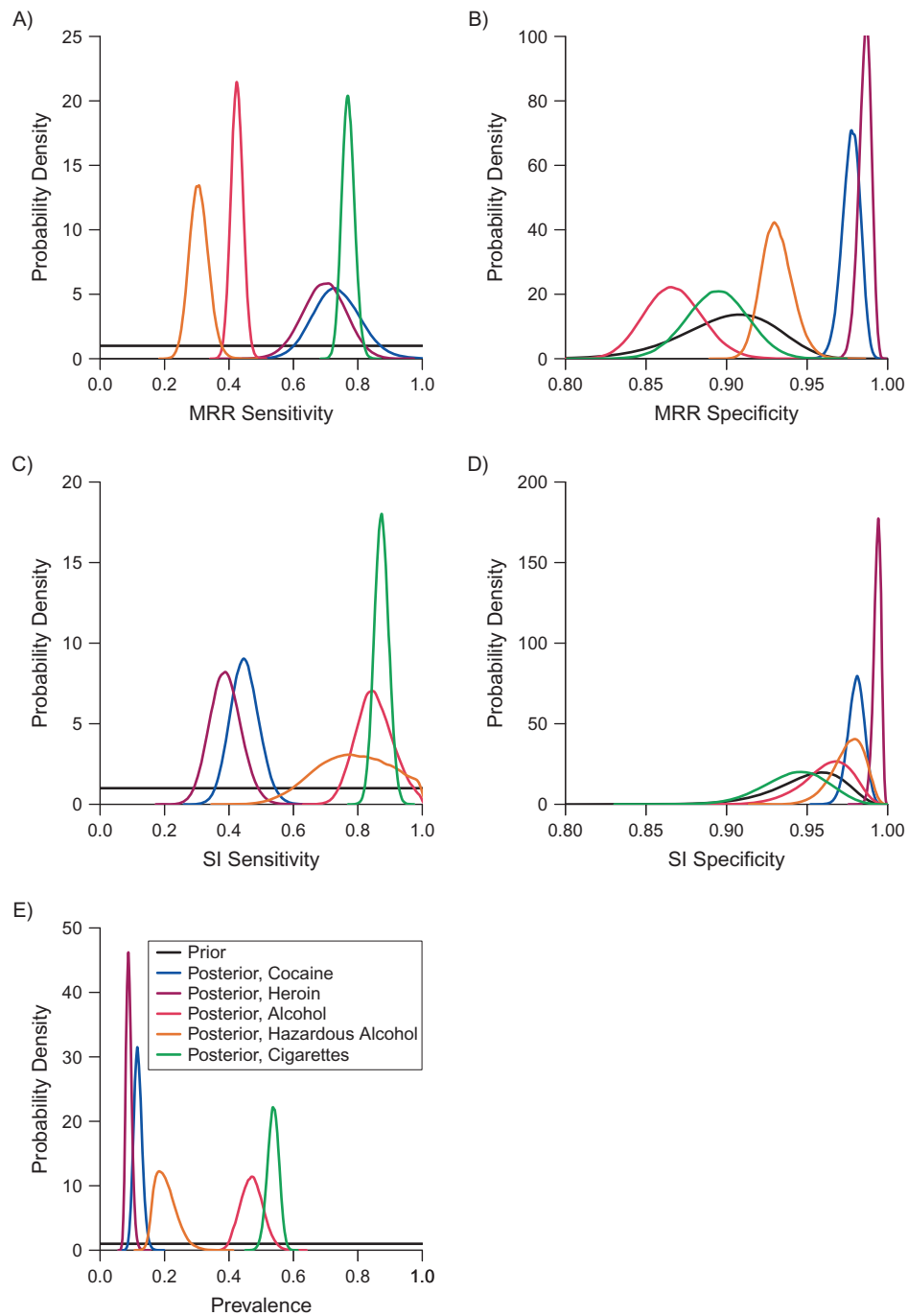


Figure 1. Prior and posterior densities for analysis of sensitivity and specificity of instruments, Johns Hopkins Human Immunodeficiency Virus Clinical Cohort, 2007–2015. A) Sensitivity of medical record review (MRR); B) specificity of MRR; C) sensitivity of self-interview (SI); D) specificity of SI; and E) prevalence of cocaine, heroin, alcohol, hazardous alcohol, and cigarette use from model 1. Prior in panels A, C, and E is distributed uniform(0, 1); prior in panel B is distributed beta(10, 90); prior in panel D is distributed beta(5, 95).

based on self-report rather than on medical record, which implies higher sensitivity of self-report, assuming specificity of both measures was very high (investigators did not directly compare the 2 measures). In contrast, we estimated higher sensitivity of MRR than SI for detecting heroin and cocaine use. Although third-party

assessments of illicit drug use are typically thought of as less sensitive than self-report (40), physician assessment of substance use for persons in regular HIV care may be unique because the physician-patient relationship is ongoing and regularly addresses sensitive topics (41). Indeed, in another setting characterized by

Table 4. Estimated Sensitivity and Specificity of Medical Record Review and Self-Interview for Detecting Recent Substance Use and Prevalence of Recent Substance Use in 2,064 Persons, Based on Different Assumptions^a, Johns Hopkins Human Immunodeficiency Virus Clinical Cohort, 2007–2015

Substance and Model	Sensitivity				Specificity				Prevalence	
	MRR		SI		MRR		SI		Median	95% CrI
	Median	95% CrI	Median	95% CrI	Median	95% CrI	Median	95% CrI		
Cocaine										
Model 0 ^b	78.4	73.5, 83.4	52.4	46.4, 58.4	100.0		100.0		13.0	11.6, 14.5
Model 1a ^c	68.8	55.7, 83.6	44.5	36.2, 53.6	97.8	96.6, 98.8	98.6	97.5, 99.5	12.5	10.0, 15.6
Model 1b	66.8	53.7, 82.4	39.9	33.0, 47.6	99.8	98.6, 100	98.9	97.6, 100	15.0	11.9, 18.6
Model 2 ^d					97.0	95.9, 98.1	98.3	97.3, 99.2	12.4	10.0, 16.4
Model 3 ^e										
3a) beta(1, 5) on pr ^e	64.9	41.8, 91.1	36.5	22.8, 50.7	97.5	96.2, 98.5	97.4	96.3, 98.5	13.0	9.1, 20.0
3b) beta(1, 7) on pr	76.0	52.6, 96.4	41.7	28.4, 53.2	97.4	96.2, 98.5	97.3	96.2, 98.4	10.9	8.2, 15.8
3c) beta(1, 3) on pr	44.0	29.2, 72.5	25.3	16.4, 41.9	97.5	96.2, 98.5	97.5	96.3, 98.6	19.5	11.8, 29.2
Heroin										
Model 0 ^b	84.9	79.5, 90.2	45.9	38.5, 53.4	100.0		100.0		8.3	7.1, 9.5
Model 1a ^c	66.8	53.7, 79.3	38.6	29.6, 48.5	98.6	97.7, 99.3	99.7	99.2, 99.9	9.3	7.7, 11.4
Model 1b	66.3	54.6, 77.7	36.1	28.4, 44.4	99.9	99.4, 100	99.9	99.4, 100	10.6	9.1, 12.5
Model 2 ^d					98.4	97.5, 99.1	99.4	98.8, 99.7	9.7	7.8, 14.5
Model 3 ^e										
3a) beta(1, 5) on pr ^e	55.7	35.6, 78.0	30.4	18.9, 44.6	98.5	97.5, 99.2	99.2	98.6, 99.7	11.1	7.8, 17.0
3b) beta(1, 7) on pr	65.6	43.7, 83.3	35.5	23.1, 47.6	98.5	97.5, 99.2	99.2	98.6, 99.7	8.4	1.7, 13.7
3c) beta(1, 3) on pr	38.9	25.8, 64.2	21.3	13.5, 36.3	98.5	97.6, 99.2	99.3	98.7, 99.7	16.1	9.9, 24.1
Alcohol										
Model 0 ^b	52.6	49.6, 55.6	81.0	78.6, 83.3	100.0		100.0		51.7	49.6, 53.9
Model 1a ^c	41.7	38.3, 45.3	84.7	74.5, 96.6	86.7	83.4, 90.4	98.6	95.7, 99.8	48.5	42.3, 55.1
Model 1b	41.8	38.4, 45.5	85.2	74.8, 95.8	86.4	83.0, 90.4	99.3	93.9, 100	48.4	42.4, 55.3
Model 2 ^d					90.9	86.9, 94.6	92.3	86.2, 97.0	51.3	44.0, 58.2
Model 3 ^e										
3a) beta(2, 4) on pr ^e	41.8	37.3, 46.2	93.6	80.6, 99.7	84.0	81.5, 86.9	95.7	91.5, 98.3	41.9	37.5, 49.0
3b) beta(2, 6) on pr	42.9	38.8, 47.3	97.2	88.5, 99.9	83.7	81.4, 86.0	94.6	89.6, 97.9	39.1	35.2, 43.5
3c) beta(2) on pr	37.5	32.4, 43.4	75.1	63.8, 91.7	85.0	81.1, 88.7	96.3	92.4, 98.7	53.3	43.4, 63.0
3d) beta(10, 90) on se _{MRR} ; beta(2) on pr	40.7	36.2, 44.9	88.5	81.5, 94.1	84.4	81.3, 87.4	96.5	92.8, 98.7	45.5	41.3, 50.4
3e) beta(5, 95) on se _{MRR} ; beta(2) on pr	41.6	37.7, 45.6	94.2	88.0, 98.0	83.9	81.5, 86.4	96.6	92.9, 98.8	42.8	39.2, 46.7
Hazardous alcohol										
Model 0 ^b	48.2	43.8, 52.6	73.3	69.4, 77.2	100.0		100.0		23.9	22.1, 25.8
Model 1a ^c	29.2	24.3, 34.7	78.4	57.1, 98.0	93.1	91.4, 95.2	99.1	97.2, 99.9	21.4	16.8, 29.4
Model 1b	29.5	24.7, 34.9	74.4	52.5, 94.1	93.8	91.9, 96.7	99.7	97.3, 100	23.1	17.9, 32.8
Model 2 ^d					95.4	93.5, 97.1	94.1	91.3, 97.3	23.0	17.1, 31.6
Model 3 ^e										
3a) beta(1, 5) on pr ^e	28.5	18.7, 37.1	84.2	56.4, 99.3	92.0	90.4, 93.3	96.7	93.8, 98.7	17.6	13.6, 26.0
3b) beta(1, 7) on pr	31.3	23.6, 39.3	92.5	70.6, 99.7	92.1	90.7, 93.3	96.3	93.2, 98.5	15.5	12.2, 20.2
3c) beta(3, 9) on pr	22.2	15.7, 31.4	65.2	48.9, 89.7	91.7	89.8, 93.3	97.3	94.6, 99.0	21.4	17.6, 30.7
3d) beta(5, 95) on se _{MRR} ; beta(1, 3) on pr	31.0	25.6, 37.5	95.0	89.5, 98.3	92.1	90.8, 93.3	97.4	94.8, 99.0	16.5	13.7, 18.9

Table continues

Table 4. Continued

Substance and Model	Sensitivity				Specificity				Prevalence	
	MRR		SI		MRR		SI		Median	95% CrI
	Median	95% CrI	Median	95% CrI	Median	95% CrI	Median	95% CrI		
Cigarettes										
Model 0 ^b	77.8	75.5, 80.1	83.3	81.2, 85.4	100.0		100.0		59.5	57.4, 61.7
Model 1a ^c	75.1	71.8, 78.9	87.2	82.9, 91.7	89.5	85.6, 93.1	97.4	92.9, 99.6	55.4	51.8, 58.4
Model 1b	76.9	72.2, 82.4	87.0	81.8, 91.9	89.6	85.1, 94.8	94.3	87.9, 99.9	54.0	48.9, 58.8
Model 2 ^d					90.4	86.4, 94.4	93.9	88.9, 97.7	54.7	50.5, 58.5
Model 3 ^e										
3a) beta(3) on pr ^e	76.3	71.9, 80.8	87.4	82.0, 92.5	88.3	84.4, 92.2	93.9	89.4, 97.4	52.9	48.7, 57.2
3b) beta(2, 3) on pr	77.1	72.7, 81.6	88.4	83.2, 93.5	87.8	83.8, 91.7	93.2	88.6, 97.1	51.6	47.2, 56.0
3c) beta(2, 3) on pr	75.2	70.2, 79.7	85.8	79.6, 90.8	89.3	85.4, 93.1	94.6	90.3, 97.9	54.9	50.7, 59.9

Abbreviations: MRR, medical record review; pr, prevalence; se, sensitivity; SI, self-interview.

^a All priors are listed in Table 1.

^b Assumes conditional independence of sensitivities and 100% specificity of both MRR and SI.

^c Assumes conditional independence of sensitivities and specificities; assumes sensitivity and specificity of MRR and SI are constant across subgroups defined by patient demographics; relaxes the assumption of perfect specificity.

^d Assumes conditional independence of sensitivities and specificities; assumes specificities of MRR and SI are constant across subgroups defined by patient demographics but allows for subgroup-specific sensitivities of MRR and SI. Because sensitivity of MRR and SI were allowed to vary by subgroup, there is not one summary measure; sensitivity according to subgroup is available in Web Table 1. By deviance information criterion, model 2 did not fit the data appreciably better than model 1.

^e Assumes conditional independence of specificities but allows for conditional dependence of sensitivities; assumes sensitivities and specificities are constant across subgroups.

frequent physician-patient interaction (women receiving prenatal care), higher prevalence of cocaine and heroin use was estimated based on medical charts than on face-to-face interview (42). We did not directly explore characteristics of the patient-provider relationship that facilitated physician awareness of substance use; other studies have suggested that patient disclosure of substance use is more likely when physicians use broad or normalizing questions (43), when assessing illicit substance use versus alcohol abuse (41), and in the context of more-frequent provider visits (44).

Our estimates of sensitivity and specificity of MRR and SI for detecting recent substance use may be transportable to other HIV clinics, given lack of evidence that sensitivity varied by patient demographics, but are not likely transportable to general practice settings. Although patients were not targeted for participation in the SI, it is possible that patients who completed the SI were more or less likely to tell the truth than patients who did not. If that was the case, our estimates of the sensitivity and specificity of the 2 measures would not be transportable.

Prevalence estimates from this study were generally comparable to those found in other studies. We estimated that 12.5% of patients in our sample had recently used cocaine. This was similar to the prevalence of illicit substance abuse or dependence estimated among persons engaged in care from 2000 to 2002 in the University of North Carolina HIV clinic, after adjusting for measurement error using a different, more resource-intensive method (45). Prevalence of self-reported alcohol use and heavy drinking was 53% and 8% in a nationally representative sample of PLWH in care in 1996 (46); we found similar prevalence of alcohol use (48.5%) but higher prevalence of hazardous alcohol use (21.4%).

Our estimate of 55.4% prevalence of current smoking is slightly higher than other estimates (typically around or greater than 40%) (47, 48). The higher prevalence of hazardous alcohol use and cigarette smoking in our sample compared with other studies may be attributable to some of the patient characteristics of our sample (largely urban, poor, and black).

Our study is not without limitations. First, models 1 and 2 assumed conditional independence of sensitivity and specificity of SI and MRR given true substance use, which may or may not be justified. SI results are not part of the medical record and not shared with the patients' providers. However, if patients who are using substances and willing to self-report substance use are also more likely to discuss their substance use with their physician, the sensitivity of SI and MRR would be correlated, and estimates from models 1 and 2 might be biased (27, 31, 49). Model 3 relaxed this assumption of conditional independence, but in the absence of additional measurements, parameters from model 3 are not uniquely identifiable and depend upon the priors placed on other parameters. Additionally, we assumed that recent substance use was the same across the 6-month reporting window for each measurement. It is possible that disagreements between the MRR and SI could have arisen from imperfectly overlapping observation windows and short-term changes in substance use. Finally, disagreements between the MRR and SI were especially pronounced for detecting hazardous alcohol use, which is likely related to the fact that documentation of substance use in the medical record is not standardized, while documentation in the SI (particularly for hazardous alcohol use) is based on standardized and validated questions.

Our study adds to the limited literature on validity of self-report of illicit drug use (40) among persons living with HIV by including a sample of patients in routine clinical care for HIV both with and without prior history of injection drug use and with and without prior history of seeking substance-abuse treatment. Furthermore, our estimates of sensitivity of SI in this population are likely more accurate than if we had used a biomarker as a gold standard. Relying on a biomarker as the gold standard may result in underestimation of the sensitivity of the SI or MRR because the biomarker will underestimate the true prevalence of substance use (50).

Assessments of time-updated substance use that rely on self-report or physician assessment alone may underestimate the true prevalence of substance use; assuming perfect specificity of both assessments generally overestimates their sensitivity. We have capitalized on the availability of 2 different measures of recent substance use to estimate their sensitivity and specificity in the absence of a gold-standard measure. Future analyses examining the effect of current substance use on HIV outcomes need to account for measurement error of substance use using a wide range of sensitivities in a sensitivity analysis.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland (Catherine R. Lesko, Richard D. Moore, Bryan Lau); Department of Epidemiology, University of North Carolina, Chapel Hill, Chapel Hill, North Carolina (Alexander P. Keil); and Department of Medicine, School of Medicine, Johns Hopkins University, Baltimore, Maryland (Richard D. Moore, Geetanjali Chander, Anthony T. Fojo, Bryan Lau).

This work was funded by the National Institutes of Health (grants U01 DA036935, U01 HL121812, U01 AA020793, U24 AA020801, and P30 AI094189).

Conflict of interest: none declared.

REFERENCES

- Lucas GM, Griswold M, Gebo KA, et al. Illicit drug use and HIV-1 disease progression: a longitudinal study in the era of highly active antiretroviral therapy. *Am J Epidemiol*. 2006; 163(5):412–420.
- Poundstone KE, Chaisson RE, Moore RD. Differences in HIV disease progression by injection drug use and by sex in the era of highly active antiretroviral therapy. *AIDS*. 2001;15(9): 1115–1123.
- Moore RD, Keruly JC, Chaisson RE. Differences in HIV disease progression by injecting drug use in HIV-infected persons in care. *J Acquir Immune Defic Syndr*. 2004;35(1):46–51.
- Hinkin CH, Hardy DJ, Mason KL, et al. Medication adherence in HIV-infected adults: effect of patient age, cognitive status, and substance abuse. *AIDS*. 2004;18(suppl 1):S19–S25.
- Baum MK, Rafie C, Lai S, et al. Crack-cocaine use accelerates HIV disease progression in a cohort of HIV-positive drug users. *J Acquir Immune Defic Syndr*. 2009;50(1):93–99.
- Arnsten JH, Demas PA, Grant RW, et al. Impact of active drug use on antiretroviral therapy adherence and viral suppression in HIV-infected drug users. *J Gen Intern Med*. 2002;17(5): 377–381.
- Samet JH, Horton NJ, Meli S, et al. Alcohol consumption and antiretroviral adherence among HIV-infected persons with alcohol problems. *Alcohol Clin Exp Res*. 2004;28(4):572–577.
- Cohen MH, Cook JA, Grey D, et al. Medically eligible women who do not use HAART: the importance of abuse, drug use, and race. *Am J Public Health*. 2004;94(7):1147–1151.
- Lucas GM, Gebo KA, Chaisson RE, et al. Longitudinal assessment of the effects of drug and alcohol abuse on HIV-1 treatment outcomes in an urban clinic. *AIDS*. 2002;16(5): 767–774.
- Cook JA, Burke-Miller JK, Cohen MH, et al. Crack cocaine, disease progression, and mortality in a multicenter cohort of HIV-1 positive women. *AIDS*. 2008;22(11):1355–1363.
- Helleberg M, Afzal S, Kronborg G, et al. Mortality attributable to smoking among HIV-1-infected individuals: a nationwide, population-based cohort study. *Clin Infect Dis*. 2013;56(5): 727–734.
- Ibañez GE, Purcell DW, Stall R, et al. Sexual risk, substance use, and psychological distress in HIV-positive gay and bisexual men who also inject drugs. *AIDS*. 2005;19(suppl 1): S49–S55.
- Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5): 615–625.
- Westreich D. Berkson's bias, selection bias, and missing data. *Epidemiology*. 2012;23(1):159–164.
- Kitahata MM, Rodriguez B, Haubrich R, et al. Cohort profile: the Centers for AIDS Research Network of Integrated Clinical Systems. *Int J Epidemiol*. 2008;37(5):948–955.
- Crane HM, Lober W, Webster E, et al. Routine collection of patient-reported outcomes in an HIV clinic setting: the first 100 patients. *Curr HIV Res*. 2007;5(1):109–118.
- Johnson T, Fendrich M. Modeling sources of self-report bias in a survey of drug use epidemiology. *Ann Epidemiol*. 2005; 15(5):381–389.
- Edwards JK, Cole SR, Westreich D. All your data are always missing: incorporating bias due to measurement error into the potential outcomes framework. *Int J Epidemiol*. 2015;44(4): 1452–1459.
- Thomas D, Stram D, Dwyer J. Exposure measurement error: influence on exposure-disease. Relationships and methods of correction. *Annu Rev Public Health*. 1993;14:69–93.
- Moore RD. Understanding the clinical and economic outcomes of HIV therapy: the Johns Hopkins HIV clinical practice cohort. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1998; 17(suppl 1):S38–S41.
- Bush K, Kivlahan DR, McDonell MB, et al. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med*. 1998;158(16): 1789–1795.
- Babor TF, Higgins-Biddle JC, Saunders JB, et al. The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care. 2nd ed. Geneva, Switzerland: Department of Mental Health and Substance Dependence, World Health Organization; 2001. http://apps.who.int/iris/bitstream/10665/67205/1/WHO_MSD_MSB_01.6a.pdf. Accessed May 8, 2017.
- National Institute on Alcohol Abuse and Alcoholism. Helping patients who drink too much: A clinician's guide. Bethesda,

- MD: National Institutes of Health; 2005. <https://pubs.niaaa.nih.gov/publications/Practitioner/CliniciansGuide2005/guide.pdf>. Accessed May 8, 2017.
24. Johnson WO, Gastwirth JL, Pearson LM. Screening without a “gold standard”: the Hui-Walter paradigm revisited. *Am J Epidemiol*. 2001;153(9):921–924.
 25. Hui SL, Walter SD. Estimating the error rates of diagnostic tests. *Biometrics*. 1980;36(1):167–171.
 26. Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *Am J Epidemiol*. 1995;141(3):263–272.
 27. Dendukuri N, Joseph L. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics*. 2001;57(1):158–167.
 28. Menten J, Boelaert M, Lesaffre E. Bayesian latent class models with conditionally dependent diagnostic tests: a case study. *Stat Med*. 2008;27(22):4469–4488.
 29. Bermingham ML, Handel IG, Glass EJ, et al. Hui and Walter’s latent-class model extended to estimate diagnostic test properties from surveillance data: a latent model for latent data. *Sci Rep*. 2015;5:11861.
 30. Hartnack S, Nathues C, Nathues H, et al. Estimating diagnostic test accuracies for *Brachyspira hyodysenteriae* accounting for the complexities of population structure in food animals. *PLoS One*. 2014;9(6):e98534.
 31. Enøe C, Georgiadis MP, Johnson WO. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev Vet Med*. 2000;45(1–2):61–81.
 32. de Araujo Pereira G, Louzada F, de Fátima Barbosa V, et al. A general latent class model for performance evaluation of diagnostic tests in the absence of a gold standard: an application to Chagas disease. *Comput Math Methods Med*. 2012;2012:487502.
 33. Plummer M, Stukalov A, Denwood M. *Bayesian Graphical Models using MCMC. R package version 46*. 2016. <https://cran.r-project.org/web/packages/rjags/rjags.pdf>. Accessed June 13, 2018.
 34. Plummer M. JAGS: A Program for the Analysis of Bayesian Graphical Models Using Gibbs Sampling. Vienna, Austria: Distributed Statistical Computing; 2003. <https://www.r-project.org/conferences/DSC-2003/Proceedings/Plummer.pdf>. Accessed June 13, 2018.
 35. Guo J, Gabry J, Goodrich B, et al. *The R Interface to Stan. R package version 2172*, 2017. <http://mc-stan.org/rstan/>. Accessed June 13, 2018.
 36. Carpenter B, Gelman A, Hoffman MD, et al. Stan: a probabilistic programming language. *J Stat Softw*. 2017;76(1):1–32.
 37. Spiegelhalter D, Best N, Carlin B, et al. Bayesian measures of model complexity and fit (with discussion). *J R Stat Soc Series B Stat Methodol*. 2002;64(4):538–639.
 38. McGinnis KA, Tate JP, Williams EC, et al. Comparison of AUDIT-C collected via electronic medical record and self-administered research survey in HIV infected and uninfected patients. *Drug Alcohol Depend*. 2016;168:196–202.
 39. Kozak MS, Mugavero MJ, Ye J, et al. Patient reported outcomes in routine care: advancing data capture for HIV cohort research. *Clin Infect Dis*. 2012;54(1):141–147.
 40. Magura S, Kang SY. Validity of self-reported drug use in high risk populations: a meta-analytical review. *Subst Use Misuse*. 1996;31(9):1131–1153.
 41. Korthuis PT, Saha S, Chander G, et al. Substance use and the quality of patient-provider communication in HIV clinics. *AIDS Behav*. 2011;15(4):832–841.
 42. O’Campo P, de Boer MA, Faden RR, et al. Discrepancies between women’s personal interview data and medical record documentation of illicit drug use, sexually transmitted diseases, and HIV infection. *Med Care*. 1992;30(10):965–971.
 43. Callon W, Beach MC, Saha S, et al. Assessing problematic substance use in HIV care: which questions elicit accurate patient disclosures? *J Gen Intern Med*. 2016;31(10):1141–1147.
 44. Korthuis PT, Josephs JS, Fleishman JA, et al. Substance abuse treatment in human immunodeficiency virus: the role of patient-provider discussions. *J Subst Abuse Treat*. 2008;35(3):294–303.
 45. Pence BW, Miller WC, Whetten K, et al. Prevalence of DSM-IV-defined mood, anxiety, and substance use disorders in an HIV clinic in the Southeastern United States. *J Acquir Immune Defic Syndr*. 2006;42(3):298–306.
 46. Galvan FH, Bing EG, Fleishman JA, et al. The prevalence of alcohol consumption and heavy drinking among people with HIV in the United States: results from the HIV Cost and Services Utilization Study. *J Stud Alcohol*. 2002;63(2):179–186.
 47. Mdodo R, Frazier EL, Dube SR, et al. Cigarette smoking prevalence among adults with HIV compared with the general adult population in the United States: cross-sectional surveys. *Ann Intern Med*. 2015;162(5):335–344.
 48. Lifson AR, Lando HA. Smoking and HIV: prevalence, health risks, and cessation strategies. *Curr HIV/AIDS Rep*. 2012;9(3):223–230.
 49. Vacek PM. The effect of conditional dependence on the evaluation of diagnostic tests. *Biometrics*. 1985;41(4):959–968.
 50. Wacholder S, Armstrong B, Hartge P. Validation studies using an alloyed gold standard. *Am J Epidemiol*. 1993;137(11):1251–1258.