HIV Incidence and Case Surveillance Branch project

Assessing the Accuracy of Self-Report of HIV Testing Behavior

Generic Information Collection Request under 0920-0840
Expiration 31 January 2013

Supporting Statement
Part B

October 31, 2011

Contact:
Jane Kelly, MD
Centers for Disease Control and Prevention
Division of HIV/AIDS Prevention
HIV Incidence and Case Surveillance Branch
Incidence and Viral Resistance Team
Phone: 404-639-6496
FAX: 404-639-8046
jkelly@cdc.gov

B. Collection of Information Employing Statistical Methods

B1. Respondent Universe and Sampling Methods

The proposed project is an observational study based on an anticipated cohort of about 900 patients in healthcare facilities in the Houston Metropolitan area. The target population of participants for the project is patients 18 years or older who receive health care from publicly-funded clinics, CTR sites, or HMO-like medical clinics that provide HIV testing or care in Houston/Harris County. Through the use of a structured questionnaire and secondary governmental databases, a facility profile of patient volume, services, and relative demography of clientele (i.e. race/ethnicity, sex) will be developed. Facility profiles will provide data regarding facilities as potential covariates with recall. Of practical concern is a facility's ability to partner with HDHHS and willingness to provide access to the clinic for recruitment and data collection activities. A two-stage sampling procedure will be conducted to obtain a representative sample. First, participating facilities will be selected from the list of all possible facilities in a manner proportional to the size of the facility. The possible facilities will include publicly-funded clinics (F1), CTR sites (F2), and HMO or similar medical clinics (F3) that provide HIV testing or care as part of their health care services in Houston/Harris County. During the second stage, participants from each facility will be asked to participate in the study by project personnel. Participants at each facility may include individuals who have never tested for HIV (G0), who have previously tested negative for HIV (G1), or who have previously tested positive for HIV (G2). Individuals will not be excluded from participation in the study at any facility based on their HIV testing statuses. The study population will be recruited from a primary care population of area residents and the sample composition of race/ethnicity is expected to be similar to demographic characteristics of each facility's neighborhood. Participants are recruited into the study by age and presence at a particular clinic during data collection. Overall numbers of participants in the three groups (i.e. G0, G1, G2) are expected to be approximately equal. Interim data analysis will be conducted to evaluate recruitment progress and appropriate adjustments will be made to facility recruitment to ensure relatively balanced groups. As each participant is enrolled, informed consent and authorization to release information is obtained (Attachment 3). Participants are eligible if they are age 18 or older and are presenting for routine health care at the selected facilities.

B2. Procedures for the Collection of Information

At the enrollment evaluation, the enrollee completes a computer-assisted self interview (CASI) using a laptop or touch-screen computer. As participants are enrolled, unique study IDs will be assigned. This unique ID will be associated with each participant throughout the life course of the study. The CASI questionnaire will collect basic participant demographics, testing history, location and results (Attachment 1). Each participant's clinical medical record will be reviewed for HIV testing histories (if any), the type and date of tests, and the test results (Attachment 2). Medical record abstraction will be limited to information five years prior to the date of interview. This timeframe limitation is necessary for practical reasons concerning availability of records. Participants whose self-report of last HIV test >5 years ago cannot be confirmed will be excluded from the analysis. The team expects one completed abstraction form and one completed questionnaire for each study participant. Medical record abstraction will serve as the "gold standard" to assess the validity and accuracy of self-reported testing histories. The targeted enrollment for participants is approximately 900 total participants.

B3. Methods to Maximize Response Rates and Deal with Non-response

Upon completion of the questionnaire, participants will receive a $10.00 token of appreciation. Participant interest is maintained by use of a short self-administered questionnaire, varying modes of data collection, remuneration and by substantial experience of the HDHHS investigators in engaging participants in research. Participants are recruited into the study by age and presence at a particular clinic during data collection. Interim data analysis will be conducted to evaluate recruitment progress and appropriate adjustments will be made to facility recruitment as needed to deal with non-response. Pre-study assessment of facility patient-care demographics, including availability of specific HIV treatment clinics or HIV testing and counseling services, at the different study sites suggests that the likelihood of recruitment stratified by HIV status and HIV testing behavior varies. Adjustments will be made to the proportion of participants recruited from the various study sites based on interim analysis in order to obtain the needed distribution of HIV positive, HIV negative and never-tested participants.

B4. Tests of Procedures or Methods to be Undertaken

The primary outcome of interest is agreement between HIV testing history indicated in the medical chart and reported by participants in the questionnaires. The overall results for this

question can be summarized in a 2x2 table and measured by the Kappa statistic, sensitivity, and specificity. Of further interest is the secondary question of how well participants recall their most recent HIV testing date and whether recall is associated with the length of time elapsed between the date of interview and the date of the testing event to be recalled or other variables. The recall accuracy will be measured by the difference in the length of time recalled and the length of time documented. These differences, if any, will be described in a scatter plot and tested for association through regression or ANOVA. A sample of the data to be collected and variables to be calculated for analysis are shown in the protocol. The agreement between participants' recalled and medically documented testing histories will be summarized by a 2x2 table and measured by the Kappa statistic and other measures such as sensitivity and specificity. Potential associations with sex, race/ethnicity, or age will be evaluated. The accuracy of recall of participants' most recent HIV test dates will be examined using calculated variables to determine the interval of time between the interview date and the recalled test date (T) and the documented test date (T_doc) and the difference between these variables (T – T_doc). The relationship between these variables will be illustrated using a scatter plot of T versus T_doc. Regression analysis and ANOVA will be used to determine if systematic recall bias exists and is associated with the length of time elapsed (T_doc), HIV testing status (G1 or G2), or other variables. The regression analysis will be conducted separately by HIV testing status (positive or negative at last HIV test) and facility type (CTR or HMO-like facility) and include other variables that potentially have significant impact on the accuracy of recalling dates. These variables include sex, age, and race/ethnicity.

B5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

Ruiguang Song, mathematical statistician at the Centers for Disease Control and Prevention (CDC) Division of HIV/AIDS Prevention (DHAP), Joseph Prejean, Team Lead for the Incidence and Viral Resistance Team (IVRT) at CDC/DHAP, and Jane Kelly, medical officer with CDC/DHAP/IVRT were consulted during the development of this protocol to ensure adequate power and appropriate sampling methods to achieve objectives for this study. To ensure a sufficient sample size for the regression analyses, the total sample size is determined by the number of subgroups. Based on CDC recommendation to include race/ethnicity as a subgroup, the total sample size for the regression analysis is

576 = [48 (HIV+) + 48 (HIV-)] x (2 facility types) x (3 race/ethnicity groups).

An equal number of observations in the never tested group is necessary. HDHHS investigators will collect the data. HDHHS investigators will perform the analysis in consultation with CDC/DHAP team. Additional staff hired to aid in collection of information will undergo extensive training in the protection of human subjects, and HIV/AIDS data confidentiality, security and safety. The study will be conducted under the auspices of the Committee for the Protection of Human Subjects at the University of Texas Health Science Center - Houston.