Attachment 4(c)

National HIV Surveillance System (NHSS)

OMB # 0920-0573

# Supplemental Surveillance Activity 1: HIV Incidence Surveillance Technical Guidance

# Technical Guidance for HIV Surveillance Programs

## HIV Incidence Surveillance

HIV Incidence and Case Surveillance Branch, Atlanta, Georgia

Cleared by HICSB Management 10/15/2012

## Notes

Cleared by HICSB Management 10/15/2012

**Table of Content**

**Content** **Page**


1. Background
   1.1 Introduction

2. Structural Requirements
   2.1 Policies and Procedures
   2.2 Definition of Reportable Information
      2.2.1 STARHS Result
      2.2.2 HIV Testing and Treatment History (TTH)
   2.3 Staffing Needs

3. Process Standards
   3.1 Collection of HIS Data Elements
      3.1.1 TTH Data Elements
      3.1.2 STARHS Result
   3.2 Entering Data into Surveillance Databases
   3.3 Creating the HIS Dataset and Transferring Data to CDC
   3.4 Data Quality
   3.5 Ensuring Security and Confidentiality
   3.6 Analyzing and Disseminating Data
   3.7 Training Staff

4. Outcome Standards

5. References

6. Appendices
   6.1. Appendix 1 National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention's Non-research Determination for HIV Incidence Surveillance
   6.2 Appendix 2 Training Resources for HIV Incidence Surveillance
   6.3 Appendix 3 Test Used by STARHS Laboratory for Recency Classification
   6.4 Appendix 4 Appendix 4.1 *STARHS Specimen Guidance: Specimens Originating from the Public Health Laboratory*
   Appendix 4.2 *STARHS Specimen Guidance: Specimens Originating from a Private/Commercial Laboratory and Sent to a Public Health Laboratory*
   Appendix 4.3 *STARHS Specimen Guidance: Specimens Originating from a Private/Commercial Laboratory and Sent Directly to the STARHS Laboratory*
   Appendix 4.4 *Guidelines for Preparing Specimens for Shipping and Transporting to the STARHS Laboratory*
   6.5 Appendix 5 Guidance for Collection and Data Entry of HIV Incidence Surveillance Information
   6.6 Appendix 6 Local HIV Incidence Estimation Guide
   6.7 Appendix 7 Completeness Report Documentation
   6.8 Appendix 8 Error Report Documentation

# Technical Guidance for HIV Surveillance Programs — Policies and Procedures for HIV Incidence Surveillance

## Background

### *Introduction*

The Centers for Disease Control and Prevention (CDC), in conjunction with a number of state and local health departments, conducts national HIV incidence surveillance (HIS), to generate annual national, state, and local estimates of the number of new HIV infections, both diagnosed and undiagnosed. HIV incidence reflects the leading edge of the epidemic, its trends, and its impact on the public's health. Using the stratified extrapolation approach[1] (SEA), CDC calculated the first national HIV incidence estimate based on a direct measure of recency of infection for 2006[2]. In 2011, using a refinement of the SEA, estimates were updated for 2006 and calculated for 2007-2009[3]. CDC expects to annually publish four year estimates based on the refined SEA in order to provide updated data on trends in incidence as the data are available.

Since 2001, CDC has funded selected state and local health departments to conduct HIS; including the collection of information on demographics, testing and treatment history (TTH), and results of a serologic marker of recent HIV infection (i.e., the serologic testing algorithm for recent HIV seroconversion [STARHS])[4] which uses a test of recent infection to classify new diagnoses of HIV infection as either recent or long-standing. In 2004, STARHS used a detuned version of a commercially available HIV diagnostic test, the Vironostika HIV-1 EIA assay, which made the test less sensitive. Because the Vironostika HIV-1 EIA had previously been approved for diagnostic purposes by the United States Food and Drug Administration (FDA), the less sensitive version of the test (Vironostika-LS) required individual consent. TTH and consent were collected during an interview with persons seeking HIV testing (pre-test) or after receiving a new HIV diagnosis (post-test). In March 2005, the FDA allowed the use of a second assay for recency testing, the BED HIV-1 Capture enzyme immunoassay (BED), for surveillance purposes only, which eliminated the need for patient consent. In 2007, CDC reduced the number of data elements required for TTH and expanded modes of collection of TTH data elements to include provider reports, medical record review, other databases and the National HIV Monitoring & Evaluation System (NHM&E), formerly Program Evaluation and Monitoring System (PEMS). Since the implementation of HIS through much of 2011, HIS data was stored in a CDC-provided Microsoft Access database. In 2011, HIS data were imported into eHARS, a document-based data system used for storing all HIV surveillance information and producing datasets for local analysis and transfer to CDC.

As an integrated component of the national HIV surveillance system, HIS incorporates into routine case reporting the collection of TTH data and recency test results in the states and cities that receive funding to conduct HIS.

The purpose of HIV incidence surveillance is to provide reliable and scientifically valid estimates of the number of newly acquired HIV infections at the local, state, and national level. CDC's human subjects protection process has determined that HIS, like other public health surveillance

activities, is not research (*See Appendix 1- National Center for HIV, STD, and TB Prevention's Non-research Determination for HIV Incidence Surveillance*).

The primary functions of HIS are to:

- Describe demographic and behavioral characteristics of newly HIV infected populations and subgroups
- Identify emerging epidemics and monitor trends in transmission
- Assist with national, state and local planning and evaluation activities
- Monitor the outcomes of national, state and local level HIV prevention programs and strategies

The tasks to achieve the functions of HIS include:

- Collaborate with CDC, laboratories, HIV testing providers and affected communities to further develop and ensure the capacity to conduct HIV incidence surveillance
- Obtain HIV testing and treatment history information on all individuals newly diagnosed with HIV infection reported to HIV Surveillance
- Collect results from tests for recent HIV infection (e.g., STARHS or other methods as they become available) necessary for the statistical estimation of HIV incidence. Currently, results are obtained through the required submission of remnant blood specimens or dried blood spot specimens from HIV diagnostic tests for recency testing using STARHS at a CDC funded laboratory. In the future, the collection of dried blood spot samples may be necessary
- Submit data monthly or as specified by the CDC using eHARS software according to data submission standards established by CDC
- Conduct systematic evaluation of HIV incidence surveillance
- Calculate and disseminate annual population-based HIV incidence estimates and promote the use of HIV incidence data for prevention and health services planning
- Ensure that the security and confidentiality procedures of the program are consistent with the requirements delineated in the *Data Security and Confidentiality Guidelines for HIV, Viral Hepatitis, Sexually Transmitted Disease, and Tuberculosis Programs: Standards to Facilitate Sharing and Use of Surveillance Data for Public Health Action*

The prerequisites (structural requirements), best practices (process standards), and outcome standards for HIS are described next.

### Structural Requirements

HIV infection may be detected at various points along the spectrum of disease, and reportable events range from reporting of HIV infection in otherwise asymptomatic persons to deaths among people with HIV. Areas conducting HIS are required to incorporate the collection of TTH variables and results from a test for recent infection from all newly diagnosed individuals aged 13 or older into HIV case reporting. TTH should be collected for all cases newly diagnosed with HIV infection, however, those diagnosed with HIV infection, Stage 3 (AIDS) are excluded from collecting results from a test for recent infection. Data should be collected in accordance with the *Technical Guidance for HIV Surveillance Programs, Vol. I: Policies and Procedures - Access to*

*Source Data and Completeness of Reporting* and the *Instructions for Completing the Data Collection Form.* HIV surveillance case report data, in combination with HIS data elements, are used to calculate population-based HIV incidence estimates.

### *Policies and Procedures*

It is important to document HIS activities and to address HIS needs in the jurisdiction's HIV surveillance policies and procedures manual to establish standardization, continuously maintain conformity of meaning for data elements, document changes over time, and develop training programs. In addition to the information listed in the *Technical Guidance for HIV Surveillance Programs, Vol. I: Policies and Procedures*, HIS specific policies and procedures should include information related to:

- Obtaining HIV testing and treatment history information on all individuals newly diagnosed with HIV infection reported to HIV surveillance
- Collecting results from tests for recent HIV infection. Currently, results are obtained through submission of remnant specimens from HIV diagnostic tests for testing at a CDC funded laboratory. Therefore, for all newly diagnosed HIV infections reported to HIV surveillance, include procedures for:
    - Establishing and maintaining collaboration and communication with public and private HIV testing laboratories (within and outside the state) to secure remnant specimens from the original diagnostic HIV test or another HIV related test conducted within 3 months of the initial diagnosis
    - Securing, transporting and tracking remnant specimens shipped to the designated laboratory for recency testing from public and private HIV testing laboratories (within and outside the state)
- Developing a STARHS eligibility list to identify which diagnostic specimens represent diagnoses of HIV infection not known to have progressed to AIDS and to inform the appropriate laboratory of the need for shipping and/or testing using an approved test for recent HIV infection
- Training all HIV case reporters (e.g. HIV testing providers, HIV care providers, etc.) and staff involved in TTH data collection (*See Appendix 2 - Training Resources for HIV Incidence Surveillance*)
- Managing data and conducting activities to ensure data quality
- Conducting systematic evaluations of HIV incidence surveillance using outcome and process standards and using evaluation results for program improvement
- Calculating and disseminating annual population-based HIV incidence estimates and promoting the use of HIV incidence data for prevention and health services planning
- Using data collected for HIS (including estimates, results of tests for recent infection and TTH data)
- Ensuring security and confidentiality procedures of the program are consistent with the requirements delineated in the *Data Security and Confidentiality Guidelines for HIV, Viral Hepatitis, Sexually Transmitted Disease, and Tuberculosis Programs: Standards to Facilitate Sharing and Use of Surveillance Data for Public Health Action*

### *Definition of Reportable Information*

### *STARHS Specimen Information and Results*

STARHS is a two test algorithm conducted on remnant specimens from HIV tests performed as part of a diagnostic algorithm or after confirmation of HIV diagnosis. The first test is used to confirm the presence of HIV specific antibodies using a standard enzyme-linked immunoassay (EIA) test and is followed by a test that determines recency of HIV infection (*See Appendix 3 - Test Used by STARHS Laboratory for Recency Classification*).

HIS coordinators in funded areas collaborate with public and private/commercial laboratories to locate, determine the disposition of, and ship remnant diagnostic blood specimens for testing using STARHS ((*See Appendix 4.1 – STARHS Specimen Guidance: Specimens Originating from the Public Health Laboratory, Appendix 4.2 - STARHS Specimen Guidance: Specimens Originating from a Private/Commercial Laboratory and Sent to a Public Health Laboratory, Appendix 4.3 - STARHS Specimen Guidance: Specimens Originating from a Private/Commercial Laboratory and Sent Directly to the STARHS Laboratory, Appendix 4.4 - Guidelines for Preparing Specimens for Shipping and Transporting to the STARHS Laboratory).*

### *Testing and Treatment History Data (TTH)*

The primary purpose of gathering HIV testing and antiretroviral (ARV) use history is to calculate a statistical weight corresponding to the probability that an individual would be tested for HIV in the STARHS recency period. This weight is used in HIV incidence estimation.

TTH data elements are used to determine testing frequency, in order to classify cases as new testers (i.e. people whose first HIV test was positive) or repeat testers (i.e. people testing HIV-positive after a previous negative HIV test). This distinction is important for incidence estimation because the probability of being classified as recent is calculated separately for new testers and repeat testers. In addition, critical TTH data elements are used to determine whether ARV use might have affected the results of the recency test.

A standard set of HIV testing and ARV use history data elements needed for the HIV incidence estimate has been developed (*See Appendix 5 - Guidance for Collection and Data Entry of HIV Incidence Surveillance Information*). This information is necessary for cases newly reported to HIV surveillance to estimate HIV incidence successfully.

### *Staffing Needs*

HIS operation requires personnel with specific skills and dedicated time to integrate HIS activities with case surveillance and other program activities (e.g. data collection, data management, data quality, data de-duplication, dissemination, etc.) to build a seamless surveillance system and maximize efficient use of resources. Generally, HIS staff should have:

- An understanding of the general principles of HIV surveillance in the jurisdiction
- Good communication skills
- Strong leadership skills
- Enthusiasm about disease reporting for public health purposes
- An orientation towards meeting HIS process and outcome standards

- Ability to work closely with CDC, other states, local sites, private providers, and laboratories

CDC recommends at least 0.5 full-time equivalent (FTE) be dedicated to the incidence surveillance coordinator (ISC) position for monitoring and evaluating the activities of HIS and 0.5 FTE for data management. Other personnel assigned to HIS may vary depending on program needs, prevalence of HIV, and available resources. Additional staff may include an epidemiologist-trainer (0.5 FTE), a laboratory liaison (0.5 FTE) and field staff for data abstraction. Finally, as HIV incidence estimation is a highly technical activity, areas funded for HIS should ensure access to a mathematical statistician. If a person with these skills is not available within the health department consultation services should be procured. Recommended staffing roles and responsibilities of HIS staff are described next.

| Position | Responsibilities | Certifications/ Expertise |
|---|---|---|
| *Incidence Surveillance Coordinator (ISC)* | • Serve as the primary point of contact for CDC on HIS<br>• Provide overall management of the HIS program and determine program goals and needs<br>• Update HIS policies and procedures manual<br>• Oversee processes for the collection of TTH and STARHS data<br>• Identify which diagnostic specimens represent new diagnoses of HIV infection not known to have progressed to AIDS<br>• Communicate with HIV case reporters and data collectors as needed regarding ascertainment and accuracy of HIS data elements<br>• Coordinate the analysis and dissemination of HIV incidence data and assess trends in new infections at the local/state level<br>• Evaluate the performance of HIV incidence surveillance using process and outcome standards<br>• Manage any employee or other service contracts related to HIS<br>• Participate in CDC site visits, trainings, and workshops<br>• Maintain the security and confidentiality of HIS data | • SAS (SAS EG) [*Desired*] |
| *Epidemiologist /Trainer* | • Serve as lead for training HIV testing providers and laboratories on HIS, including development/modification of surveillance area-specific training materials<br>• Verify the adherence to TTH data collection standards and proper submission of TTH data elements<br>• Conduct quality control activities to address the accuracy of the data collected<br>• Provide feedback to HIV case reporters and data collectors on TTH completeness and timeliness<br>• Coordinate the collection of TTH data elements from HIV case reporters and data collectors<br>• Use process and outcome standard evaluation results for program improvement<br>• Provide consultation and technical assistance on TTH data collection<br>• Participate in CDC site visits, trainings, and workshops as | • Training<br>• Data collection<br>• Data quality<br>• SAS (SAS-EG) |

| | | |
|---|---|---|
| | appropriate<br>• Maintain the security and confidentiality of HIS data | |
| *Laboratory Liaison* | • Coordinate with public and private HIV testing laboratories (within and outside the state) to secure remnant specimens from the original diagnostic HIV test or other HIV related test conducted within 3 months of the initial diagnosis<br>• Coordinate/communicate with public and private HIV testing laboratories (within and outside the state) to arrange transport of remnant specimens to the designated laboratory for recency testing<br>• Maintain communications between public health, private, and community laboratories, the STARHS laboratory and the ISC<br>• Establish specimen tracking mechanisms<br>• Serve as the subject matter expert on the preparation and shipping of specimens to the STARHS laboratory including quality control procedures<br>• Oversee specimen identification and tracking<br>• Provide feedback to laboratories on specimen quality and quantity as well as frequency of specimen shipments<br>• Participate in CDC site visits, trainings, and workshops as appropriate<br>• Maintain the security and confidentiality of HIS data | • Air Transport Association (IATA) Training/Specimen handling/aliquotting experience<br>• Laboratory database experience<br>• Encryption software |
| *Data Manager* | • Oversee processes for data entry of incidence data elements<br>• Conduct data quality assessments, dataset creation and transfers, and data management programs<br>• Maintain documentation of HIS data management processes in policies and procedures manual<br>• Participate in process and outcome standard evaluations<br>• Serve as subject matter expert on HIV incidence data elements, data set preparation for local and national incidence estimation, and daily data management<br>• Participate in CDC site visits, trainings, and workshops<br>• Maintain the security and confidentiality of HIS data | • SAS (SAS EG)<br>• Encryption software |

## Process Standards

HIS involves the following processes:

- Ongoing, collect HIV testing and treatment history information on all individuals aged 13 years or older newly diagnosed with HIV infection and reported to HIV surveillance
- Ongoing, collect results from tests for recent HIV infection for all individuals 13 years and older newly diagnosed with HIV, excluding those with HIV infection, Stage 3 (AIDS). Obtain results through required submission of a remnant specimen from an HIV diagnostic test to a CDC funded laboratory as needed
- At least monthly, enter HIS data into the eHARS database (version 3.3 or later)
- Monthly, transfer HIS data electronically to CDC using eHARS software according to data submission standards established by CDC
- Routinely, conduct data quality control activities including data error resolution and monitoring monthly data completeness

- Calculate and disseminate annual population-based HIV incidence estimates and promote the use of HIV incidence data for prevention and health services planning
- At least annually, conduct a systematic evaluation of HIV incidence surveillance using outcome and process standards and use evaluation results for program improvement
- Ongoing, ensure that data handling procedures comply with Security and Confidentiality Guidelines (See *Data Security and Confidentiality Guidelines for HIV, Viral Hepatitis, Sexually Transmitted Disease, and Tuberculosis Programs: Standards to Facilitate Sharing and Use of Surveillance Data for Public Health Action* and *Model State Public Health Privacy Act*)
- At least annually, provide training to HIV reporters and data collectors in HIV surveillance and HIS methods

### *Collection of HIS Data Elements*

As a general principle, TTH data should be collected for all individuals aged 13 years or older newly reported to the HIV surveillance system by all providers of HIV testing. Results from tests for recent HIV infection should be collected for all individuals 13 years and older newly diagnosed with HIV, excluding those with HIV infection, Stage 3 (AIDS).

### *Cases Out of Jurisdiction*

HIS funded jurisdictions should collect TTH data (including ARV use variables) and STARHS results for all persons newly reported to HIV surveillance and work with other areas to exchange needed data. Refer to the *Technical Guidance for Surveillance Programs, Vol. I: Case Residency* for information about interstate/reciprocal notification of reportable diseases. For cases reported to an HIS funded jurisdiction and residing in any HIS funded jurisdiction HIS areas are expected to collect TTH data and ensure that a remnant of the diagnostic specimen is tested using STARHS for all cases (not known to be AIDS – Stage 3). All TTH and STARHS data should be forwarded to the state of residence at HIV diagnosis if the state is funded for HIS. In addition, ARV use data typically collected through HIS are critical to Molecular HIV Surveillance (MHS; formerly variant, atypical, and resistant HIV surveillance) data analyses; HIS areas should collaborate with other jurisdictions to provide needed data.

### *STARHS Specimen Information and Results*

Collecting results from tests for recent HIV infection for all individuals 13 years and older newly diagnosed with HIV, excluding those with HIV infection, Stage 3 (AIDS) is required for HIS. For all newly diagnosed HIV infections reported to HIV surveillance, funded health departments are required to:

- Identify and locate diagnostic specimens that represent diagnoses of HIV infection not known to have progressed to AIDS at the time of HIV diagnosis
- Secure remnant specimens from the original diagnostic HIV test or other HIV related test conducted within 3 months of the initial diagnosis from public and private HIV testing laboratories (within and outside the state).
- Determine the disposition of specimens for recency testing using STARHS by checking the quarterly specimen inventory list sent by the STARHS laboratory and by developing a monthly STARHS eligibility list for both the public health laboratory and the STARHS

laboratory
- Establish mechanisms to track shipments as appropriate and receipt of results
- Establish regular communication between the ISC, HIS staff, the public health laboratory, private and commercial laboratories, and the STARHS laboratory

Results from tests for recent infection are obtained through locating and determining the disposition of specimens, preparing STARHS eligibility lists, and transporting remnant HIV-positive specimens to the STARHS laboratory. This laboratory is a CDC funded centralized laboratory responsible for receiving and storing specimens, conducting tests for recent infection within STARHS, and communicating results of tests for recent infection to jurisdictions conducting HIV incidence surveillance

Laboratories that perform routine diagnostic confirmatory/supplemental HIV testing and immunological status tests such as CD4, or viral load counts should report results to the state/local health department surveillance program per existing requirements. For the purposes of this guidance, laboratories that are sources of remnant specimens for recency testing have been classified into 3 types:

1) Commercial laboratories that process specimens from many states and/or jurisdictions (included in this category are: Quest Diagnostics Inc., Laboratory Corporation of America [LabCorp], ARUP Laboratories, , and Mayo Clinic)
2) Private laboratories, which include smaller private/university/hospital or medical center laboratories that provide service primarily at the state or local level
3) State and/or local public health laboratories that work in collaboration with other arms of the nation's public health system, to provide clinical diagnostic testing, disease surveillance, environmental and radiological testing, emergency response support, applied research, laboratory training and other essential services to the communities they serve.

In each surveillance area, all laboratories that perform HIV diagnostic tests for residents of the surveillance area should be identified from a review of local HIV surveillance data and state laboratory licensing records. An updated directory of laboratory contacts at all reporting laboratories should be maintained to facilitate communication in the event that reporting or shipping of specimens is disrupted or that changes in policy or procedures need to be communicated. The laboratories identified should be contacted to request that remnants of HIV-positive diagnostic specimens be made available for recency testing using STARHS. The reports that these laboratories send to the state/local health department surveillance programs per the established reporting requirements should include the specimen accession numbers.

Sending specimens from a laboratory that performs the initial screening HIV EIA test (originating laboratory) to a second laboratory for confirmatory/supplemental HIV testing (reference laboratory) is a common practice. A specimen may be assigned a different accession number at each laboratory that receives it; therefore, it is important to ensure that the accession number that is assigned by the laboratory that ships the specimen to the STARHS laboratory is transmitted to the surveillance program as part of the laboratory reporting process.

HIS staff, using routine HIV surveillance reporting procedures, determines whether the specimen represents the person's first reported positive HIV test result in the HIS area or was from an HIV related test drawn within three months of the initial reported positive HIV test result and is eligible for recency testing.

A specimen should be tested for recency of HIV infection if

- the specimen represents the diagnostic specimen (the HIV-positive specimen that led, or should have led an adolescent/adult HIV case to be reported to HIV surveillance) or
- the diagnostic specimen is unavailable and the specimen was drawn within 3 months of the diagnostic specimen and
- the specimen was drawn for an HIV-related test (including viral load, polymerase chain reaction [PCR] test, and CD4 count)

A specimen should not be tested for recency of HIV infection if:

- the individual had a previous specimen that was tested for recency
- the specimen was drawn more than 3 months after the diagnostic specimen
- the individual is known to have HIV infection, Stage 3 (AIDS) at diagnosis
- the individual has a positive HIV-1 viral load or qualitative RNA specimen, but a negative HIV-1 antibody test result on or after the collection date of the positive HIV-1 viral load
- the individual has a positive rapid test and an undetectable viral load and no other diagnostic test results are available to confirm/supplement diagnosis
- the individual has a HIV-2 antibody positive diagnosis or is positive for both HIV-1 and HIV-2 antibodies (dual infection)

Specimens will be classified as HIV-positive at one of three types of laboratories previously described. For each positive HIV test reported to an HIS jurisdiction the disposition of that specimen as either eligible or ineligible to be tested for recency of HIV infection must be determined. At the public health laboratory only specimens deemed eligible to be tested for recency of HIV infection are shipped to the STARHS laboratory whether specimens were tested at the public health laboratory or shipped there by another diagnostic laboratory. The disposition of the specimens must be communicated by HIS staff to the public health laboratory (*See Appendix 4.1 – STARHS Specimen Guidance: Specimens Originating from the Public Health Laboratory, Appendix 4.2 - STARHS Specimen Guidance: Specimens Originating from a Private/Commercial Laboratory and Sent to a Public Health Laboratory* ).

Commercial and private laboratories may ship all remnant HIV-positive specimens directly to the STARHS laboratory. The disposition of these specimens must be communicated by HIS staff to the STARHS laboratory. Communication with the STARHS laboratory related to the disposition of remnant specimens is accomplished using the STARHS Eligibility List (*See Appendix 4.3 - STARHS Specimen Guidance: Specimens Originating from a Private/Commercial Laboratory and Sent Directly to the STARHS Laboratory*). For instructions for the shipment of remnant HIV-positive diagnostic specimens please see *Appendix 4.4 - Guidelines for Preparing Specimens for Shipping and Transporting to the STARHS Laboratory*.

As future tests of recent HIV infection become available, programs must adapt to CDC guidance for the collection of results from these tests.

### *Testing and Treatment History (TTH)*

The general principle in TTH data collection is to obtain the most accurate information available to characterize a person's HIV testing and treatment history. Critical TTH data elements include standard variables used for HIV incidence estimation[1, 3] and HIS evaluation.

TTH data are reported by providers and surveillance staff using the area standard reporting procedures or other procedures that meet the routine security and confidentiality guidelines for HIV surveillance. These data:

- Should be included for all individuals aged 13 years or older that are newly reported to the HIV surveillance system by all individuals responsible for reporting diagnoses of HIV infection
- May be collected from multiple sources of information, including patient interview, medical records, laboratory reports, and other databases including NHM&E
- Should be collected following the tenets of document-based surveillance
- May be collected when investigating a case, completing a case report form, or conducting interviews as part of counseling and testing services, partner services or other means

TTH fields are included in the revised CDC adult case report form (ACRF) (*See Appendix 5 - Guidance for Collection and Data Entry of HIV Incidence Surveillance Information* and *Technical Guidance for HIV Surveillance Programs, Vol. II: Instructions for Completing the Data Collection Form*). It is important to examine all information available and ensure that data collected most accurately reflect the patient's actual HIV testing history and antiretroviral use.
CDC has developed materials for use in training HIV reporters and data collectors to obtain HIV testing and treatment history data. (*See Appendix 2 - Training Resources for HIV Incidence Surveillance*)

### Entering Data into Surveillance Databases

HIS data may be entered manually or imported into eHARS, but should follow the principles of document-based surveillance. TTH data are entered into an eHARS TTH document; separate eHARS documents are used to enter and store data from different forms, or from multiple sources, or to update information. This important feature allows CDC and the project area to track the relative proportions, and data quality and completeness of TTH documents collected through various reporting methods (e.g., patient interview, medical record review, or passive case reporting) for the purpose of program improvement. Data entry staff is expected to enter data in eHARS as reported on the data collection form. If conflicts within a data collection form are identified, the HIS program should have procedures in place to resolve the data conflicts. If the conflict is the result of a typographical error in a single document, the error should be corrected in that document.

STARHS eligibility lists prepared by HIS staff are sent either to the public health laboratory (which will assign STARHS IDs to eligible specimens and ship them to the STARHS laboratory)

or to the STARHS laboratory for specimens sent directly from private or commercial laboratories (which ship specimens identified only by specimen ID or accession number). The STARHS laboratory and the PHL will send back to the HIS staff a link file for eligible specimens matching STARHS ID numbers assigned by the appropriate laboratory to corresponding specimen accession numbers for specimens that were identified as eligible by HIS staff. Following testing for recent infection, the STARHS laboratory will send results (identified only by STARHS ID) to HIS staff.

There are two approaches for entering specimen and STARHS information into eHARS based on whether or not the jurisdiction uses an external data tracking system to manage STARHS specimen data. The first approach applies to jurisdictions that manually enter or import STARHS specimen data into eHARS before importing STARHS results. STARHS specimen required and recommended information for eligible specimens should be entered into eHARS on a regular basis, including STARHS ID, first, middle and last name of the case reported, specimen collection date, STARHS test type, specimen type, relevant specimen IDs or accession numbers, and if necessary information on the reasons specimens were not sent for testing for recent infection (e.g. quantity not sufficient, etc.). It is important to note that before importing STARHS result files, the STARHS specimen data, particularly the STARHS ID, collection date and test type must be entered into eHARS as a Lab document. In this approach, when STARHS results are returned, they will be matched to cases in eHARS using the previously entered/imported STARHS ID. Some important STARHS specimen information including specimen collection date, accession number, specimen ID and the case's first, middle and last names will be extracted (using the CDC provided SAS program) from the previously entered STARHS specimen lab document and included in the STARHS result import file. After importing the STARHS result file, a new Lab document containing the results and specimen collection date, accession number, specimen ID and the case's first, middle and last names will be added to each corresponding case. Thus, each case will have at least two documents associated with a given recency testing event. For analysis, the data from all Lab documents for a unique STARHS ID should be combined.

The second approach applies to jurisdictions that manage STARHS specimen information using a locally designed data tracking system outside of eHARS. Approach 2 combines the locally-stored specimen data (from the tracking database) with the STARHS results (from the STARHS lab) outside of eHARS and allows the import of both into eHARS on a single document. The external data tracking system should include the following required and recommended information: STATENO, first name, middle name, last name, STARHS ID, laboratory specimen IDs or accession numbers, specimen collection date, and specimen type. To combine STARHS results with the specimen data, one needs to first extract the specimen data from the external tracking system as a CSV file (see Table 1 which provides the structure of the CSV file along with variable names, descriptions, types and acceptable values), then run the SAS import preparation program provided by CDC. The SAS import preparation program combines the specimen data and STARHS result data by STARHS ID and prepares an auto-import STARHS data file based on STATENO. Importing the auto-import STARHS data file into eHARS will add a STARHS lab document containing both specimen and STARHS result data for a given specimen to existing eHARS cases. If the jurisdiction uses the CDC-provided SAS program, the eHARS auto-import file will contain only specimen data for those specimens that have a matching STARHS ID in the

STARHS result file. The specimen data without a matching STARHSID in the STARHS result file are likely to be specimens that are waiting for STARHS results, and therefore will be matched to future STARHS results. Most surveillance areas also maintain in their tracking system data on specimens that are not sent for testing for recent infection. Information on the reason specimens were not sent for testing for recent infection cannot be imported into eHARS, therefore data for these specimens should be manually entered into eHARS, along with required or recommended specimen information such as specimen collection date, STARHS test type, specimen type and relevant specimen IDs. With Approach 2, surveillance staff can enter data into eHARS manually when necessary.

Table 1: The STARHS specimen data CSV file structure with variable name, label, type, and acceptable values. Note: It is REQUIRED to maintain the order of the variables; the header row is not necessary.

| STATENO | FIRST | MIDDLE | LAST | SSTARHSID | SSTATEID | LSrceID | LDteObt | LSpecTy |
|---|---|---|---|---|---|---|---|---|
| State No | First name | Middle name | Last name | STARHS specimen ID | State Lab or other specimen ID | Source Lab specimen ID | Specimen collection date | Specimen Type |
| Char(35) | Char(50) | Char(50) | Char(50) | Char(15) | Char(50) | Char(50) | mm/dd/yyyy | Char(10) |
| | | | | | | | | URN<br>BLD<br>OTH<br>SAL<br>UNK |

A list of data elements necessary for HIS can be found in the *Local HIV Incidence Estimation Guide – Appendix 6*. Information regarding HIS data entry and variables that are required or recommended for eHARS HIS documents are designated in the *Guidance for Collection and Data Entry of HIV Incidence Surveillance Information - Appendix 5*.

### *Creating the HIS Dataset and Transferring Data to CDC*

eHARS is the only data entry and management system that state or local health departments should use for HIS. HIS data collected from other sources, including the excel spreadsheet from the STARHS laboratory that contain STARHS results identified by STARHS ID, National HIV Monitoring and Evaluation system (NHME), formerly known as the Program Evaluation Monitoring System (PEMS) or other local databases should be manually entered or imported into eHARS.

Jurisdictions should update and export eHARS person-based and document-based datasets and run the CDC-provided SAS MACRO (HIS Data Export Program – provided by the CDC HIS Program Coordinator upon request) to generate an incidence SAS dataset containing all cases reported by the jurisdiction with accompanying incidence data (TTH and STARHS result). Until National Data Processing (NDP) of eHARS data is implemented, jurisdictions should submit the data to CDC via secure electronic data methods between the 1st and 15th of each month. After the implementation of NDP, jurisdictions should submit the previously described dataset to CDC via secure electronic data methods between the 1st and 15th of the first month of each quarter. Data transmitted

to CDC must not include personal identifiers and must be encrypted and password protected see the *Data Security and Confidentiality Guidelines for HIV, Viral Hepatitis, Sexually Transmitted Disease, and Tuberculosis Programs: Standards to Facilitate Sharing and Use of Surveillance Data for Public Health Action*.

### *Data Quality*

CDC has developed programs that generate reports to be used to improve the quality of incidence data. There are two data quality measures: the HIV incidence completeness and error reports.

The HIV incidence completeness report provides a snapshot of the proportion of HIV cases diagnosed within a three year period that have required TTH data elements and STARHS result. Completeness is based on all HIV cases diagnosed in a specific year. Percentages are based on cases reported from each HIS area and residing in any area funded for HIS. The HIV incidence completeness report serves as a data quality monitoring tool to measure HIS data completeness, identify problems with data transfer and monitor progress toward meeting target performance levels and data quality standards. High completeness will support the HIV incidence surveillance programs to meet outcome standards. If issues are identified, CDC will provide technical assistance to improve the completeness of data.

The HIV incidence error report monitors data entry errors and data inconsistencies between variables. It is based on cases reported from an HIS area and residing in areas funded for HIS. The validity and accuracy of the HIV incidence estimates are limited if the data contain errors or a large proportion of information is missing or unknown. If errors and inconsistencies are checked and fixed routinely, programs can detect problems with data entry and data collection, and improve data quality.

Monthly completeness reports and quarterly error reports serve as data quality monitoring tools for CDC epidemiologists and state/local HIS staff to use in understanding HIS data and in working together to improve the validity and accuracy of the data. HIS sites investigate and determine the resolution for errors based on the recommendations provided by CDC (*See Appendix 7- Completeness Report Documentation; Appendix 8 - Error Report Documentation*).

### *Ensuring Security and Confidentiality*

Data collected for HIV incidence surveillance are considered part of routine surveillance and should be held to the standards of security and confidentiality for HIV surveillance outlined in the *Data Security and Confidentiality Guidelines for HIV, Viral Hepatitis, Sexually Transmitted Disease, and Tuberculosis Programs: Standards to Facilitate Sharing and Use of Surveillance Data for Public Health Action*. Policies and procedures, based on the guidelines and local laws, are already in place at state and local health departments and are used to secure hard copies and electronic data to protect the confidentiality of persons reported as having HIV infection. Specimens and associated information that are carried through US mail or carrier services should be treated as case information and should follow all security and confidentiality requirements that apply to physical and data security. Access by all staff to information in eHARS, HIV testing and ARV use history, and STARHS data is governed by the same security and confidentiality

requirements. STARHS results using the BED are intended for HIV surveillance purposes only to support population studies and cannot be interpreted at the individual level for clinical decisions.

### *Analyzing and Disseminating HIV Incidence Estimates*

Population-based HIV incidence is estimated using a statistical approach that is analogous to that used to estimate a population total from a sample survey. Following the sample survey approach, the number of new HIV infections in the population is estimated from a sample composed of all newly diagnosed HIV infections in the selected period of time classified as "recent" using the BED by assigning a sampling weight to each result in the sample. This weight is the inverse of the estimated probability that a seroconverted person would be identified as having a recent infection based on the test for recent infection. Data from persons who choose to have a confidential HIV test and who test positive are used to estimate the incidence of HIV infection, both diagnosed and undiagnosed nationally and in participating areas. The method used to generate the population-based incidence estimate was initially described by Karon et al in *Estimating HIV Incidence in the United States from HIV/AIDS Surveillance Data and Biomarker HIV Test Results*[1]

HIV prevention programs can use HIV incidence estimates for targeting resources, and monitoring and evaluating prevention activities locally, regionally, and nationally. CDC has the primary responsibility for analyzing interpreting, and disseminating national HIV incidence estimates using data from HIS areas as appropriate. Results will be presented at conferences and published in peer-reviewed journals. The number of representative authors from participating areas and CDC will be determined for each presentation or paper. State or local HIS areas should consider developing their own HIV incidence estimates following *Appendix 6 - Local HIV Incidence Estimation Guide*.

### *Training Staff*

In accordance with the *Data Security and Confidentiality Guidelines for HIV, Viral Hepatitis, Sexually Transmitted Disease, and Tuberculosis Programs: Standards to Facilitate Sharing and Use of Surveillance Data for Public Health Action*, HIS staff must also receive training on security and confidentiality procedures, and must sign a confidentiality statement upon being hired and annually thereafter. Because HIS is a fully integrated component of HIV surveillance, all HIS staff should receive training in the local policies and procedures for case surveillance including:

- Active and passive surveillance methods
- Laboratory reporting mechanisms
- Data management processes

Besides the general training mentioned above, the HIS laboratory liaison, data manager and ISC/ epidemiologist should receive specific job related trainings such as:

Laboratory Liaison:

- Role of Public Health Laboratory in incidence surveillance
- Specimen handling i.e. specimen numbering, storage and retention
- Aliquotting and specimens shipment to the STARHS laboratory according to guidelines
- Entering data and generating reports for specimen tracking and determining specimen

eligibility for recency testing using STARHS

Data Manager:

- HIS data entry and data quality assessment
- Secure electronic data transport procedures
- Specimen eligibility and, tracking and maintaining results (electronic and hard copies)
- Creating reports for HIV care and prevention programs

ISC/Epidemiologist:

- Incidence estimation training
- Program monitoring and evaluation
- eHARS training

## Outcome Standards

Outcome standards described in the "Introduction to Policies and Procedures" and "Data Quality" sections of *Technical Guidance for HIV Surveillance Programs, Vol. I: Policies and Procedures* can be applied to HIS. These sections address completeness, case ascertainment, timeliness of reporting, evaluation of standard data edits, and missing/unknown information. Meeting case surveillance standards for case ascertainment and timeliness is essential for HIS to be successful given the time-sensitive nature of HIS data elements, including TTH data and STARHS result. The quality of the HIV incidence estimate depends on the quality and completeness of data included in the HIS system. All outcome standards for HIS relate only to cases that reside within the surveillance area at the time of diagnosis.

Data Quality

- 97% of case records pass all selected data edits related to HIS data, measured at 12 months after the close of the report year, with a target of 100%

Completeness of Testing and Treatment History Data

- At least 85% of newly diagnosed HIV disease cases reported for a calendar year should have HIV testing and treatment data assessed at 12 months after the end of the diagnosis year

Completeness of STARHS Result

- At least 60% of new diagnoses of HIV infection, excluding cases diagnosed with AIDS within 6 months of HIV diagnosis, reported for a calendar year have a STARHS result from a specimen obtained at, or within 3 months of HIV diagnosis assessed at 12 months after the end of the diagnosis year

## References


1. Karon, J.M., Song, R., Brookmeyer, R., Kaplan, E.H., & Hall, H.I. (2008). Estimating HIV incidence in the United States from HIV/AIDS surveillance data and biomarker HIV test results. Statistics in Medicine 27(23): 4617-4633
2. Hall, H.I., Song, R., Rhodes, P., Prejean, J., An, Q., et al. (2008). Estimation of HIV incidence in the United States. Journal of the American Medical Association 300(5): 520-529.


Cleared by HICSB Management 10/15/2012

3. Prejean J, Song R, Hernandez A, Ziebell R, Green T, et al. (2011) Estimated HIV Incidence in the United States, 2006–2009 PLoS ONE August 2011 6 (8).
4. Janssen RS, Satten GA, Stramer SL, et al. New testing strategy to detect early HIV-1 infection for use in incidence estimates and for clinical and prevention purposes. Janssen RS, Satten GA, Stramer SL, et al. New testing strategy to detect early HIV-1 infection for use in incidence estimates and for clinical and prevention purposes. *JAMA* 1998; 280:42–48.
5. Parekh.BS, Hanson DL, Hargrove J, Branson B, Green T, Dobbs T, et al,(2010). Determination of Mean Recency Period for Estimation of HIV Type 1 Incidence with the IgG capture BED enzyme immunoassay (BED-CEIA) in Persons Infected with Diverse Subtypes. AIDS Research and Human Retroviruses 26(00): 1-9

# National Center for HIV, STD, and TB Prevention's Non-research Determination for HIV Incidence Surveillance

## NCHSTP Research/Non-research Determination

### (Request to Classify Project as Not Involving Human Subjects or Research)

This form should be used to submit to NCHSTP ADS materials for projects involving CDC investigations that are not subject to human subjects regulations. Projects are eligible for this classification either as "non-research" projects (primary intent is not to generate generalizable knowledge) or as research projects that do not involve identifiable human subjects. Such projects do not require submission to the CDC Human Subjects Office for IRB review. Do **NOT** use this form for IRB "EXEMPT" research.

Project Title___ Standard HIV Incidence Surveillance Procedures Guidance _____

Project Locations/Sites: _____33 funded incidence sites, eventually all HIV/AIDS Surveillance areas___

Project Officer(s)____Lisa M. Lee, PhD __ Division:__ DHAP-SE _____ Telephone: __404.639.2052

Proposed Project Dates: Start: 03/__01__/_2005_ End: N/A, routine surveillance _____/_____/_____

Categories of data collection that do not constitute human subjects research include are listed below. Please check appropriate category:

_X__ **I. Activity is not research** . Primary intent is public health practice or a disease control activity.
- ___ **A.** Epidemic/endemic **disease control** activity; collected data directly relate to disease control needs.
- _X__ **B.** Routine **disease surveillance** activity; data used for disease control program or policy purposes.
- ___ **C. Program evaluation** activity; data are used primarily for that purpose.
- ___ **D. Post-marketing surveillance** of efficacy and/or adverse effects of a new regimen, drug or device.
- ___ E. **Activity is purely administrative** (e.g., purchase orders or contracts for services or equipment) and not related to research [this category I-E may be determined by Divisional ADS]

-OR-

___ **II. Activity is research but does NOT involve identifiable human subjects.**
- ___ **A.** Activity is research involving collection/analysis of data about health facilities or other organizations or units which are not individual persons....**or**...
- ___ **B.** Activity is research involving data and/or specimens from deceased persons...**or**...
- ___ **C.** Activity is research using unlinked anonymous data or specimens: **All** (1-4) of the following are required:
  - ___ 1. No contact with human subjects is involved for the proposed activity...**and**...
  - ___ 2. Data or specimens are/were collected for another purpose...**and**...
  - ___ 3. No extra data/specimens are/were collected for **this** purpose...**and**...
  - ___ 4. Identifying information either was not obtained **or** has been removed so that data cannot be linked or re-linked with identifiable human subjects. (Note: under certain conditions, research *may* qualify as non-human subjects when identifiers are removed by local staff; contact NCHSTP ADS office for details.)

Attach project description (standard format at end of this form) in enough detail to clarify its non-human subject research nature. Submit through division ADS/Director to: NCHSTP ADS, Attn: Janella Dodson (MS E-07)

☐ Check here if this request is an **amendment** of an existing non-research determination.

Approval initials: _____ 3/2/05 _____ 3/2/05
**Branch/Section Chief** Date **ADS or Div. Director** Date

*res-nonres.det.form.0704*

**Page 2**

Project Title_____ Standard HIV Incidence Surveillance Procedures Guidance _____

**NCHSTP ADS Review** **Date rec'd in NCHSTP ADS Office:**_____

~~X~~ **Concur, project does not constitute human subjects research** This project is national HIV/AIDS surveillance + not research.

**or**

___**Project constitutes human subjects research**, submission for Human Subjects review required

**Comments/Rationale:**

**Additional Comments:**

1. This form cannot be used to document "IRB Exempt Research," which must instead be submitted to the CDC IRB. (Please contact the NCHSTP ADS Office for details).
2. Although CDC Human Subjects (IRB) review is not required in this instance, investigators/project officers are expected to adhere to ethical principles and standards by respecting and protecting to the maximum extent possible the privacy, confidentiality and autonomy of participants. All applicable State and Federal privacy laws must be followed.
3. Although this project does not constitute human subjects research, informed consent may be appropriate. Information disclosed in the consent process should address the eight standard consent elements.
4. Other:

**Signed:** _____ 3/9/05

Terence Chorba, MD, MPH, MPA, MA
Acting Associate Director for Science, NCHSTP
National Center for HIV, STD, and TB Prevention

Date

*res-nonres.det.form.0704*

## Appendix 2 - Training Resources for HIV Incidence Surveillance

The **TTH Training Materials (Job aids)** were developed to assist HIV Incidence Surveillance sites in training staff to gather, record and enter testing and treatment history (TTH) data, and to inform health care providers about their role in HIV incidence surveillance. In addition to the TTH Job Aids, PowerPoint slides were presented at the 2010 HIV Incidence Surveillance Workshop: Data Quality and Estimation and are available on on the HICSB SharePoint at **https://partner.cdc.gov/sites/NCHHSTP/HICSB/default.aspx**to assist with training efforts.

The **TTH Job aids** have been developed for four different audiences:

1. **TTH Training Materials for Health Care Provider (English and Spanish)**
   Intended for health care providers who care for HIV patients and record information in the medical chart or on a provider case report form
   - **Pocket card (4x6") –** reminder of key TTH questions
   - **Quad-fold brochure -** basic information on HIV incidence surveillance and the critical role health care providers play in gathering accurate and complete TTH information

2. **TTH Training Materials for Patient Interview**
   Intended for personnel who interview patients directly for TTH
   - **Patient Interview bullet point reference -** designed to carry on a clipboard as a quick reference when recording TTH from patient interviews
   - **Patient Interview *TTH ACRF*** - designed to post on a workspace bulletin board

3. **TTH Training Materials for Medical Record Review**
   Intended for personnel who perform medical records abstraction to obtain TTH information
   - **Medical Record Review bullet-point reference -** designed to carry on a clipboard as a quick reference during medical record abstractions
   - **Medical Record Review *TTH ACRF*** - designed to post on a workspace bulletin board

4. **TTH Training Materials for Data Entry**
   Intended for personnel who enter TTH information into an electronic database
   - **Data Entry bullet-point reference -** quick reference for use at the data entry station
   - **Data Entry *TTH ACRF* -** designed to post on a workspace bulletin board

All files are available on the HICSB SharePoint at **https://partner.cdc.gov/sites/NCHHSTP/HICSB/default.aspx** as modifiable files so that sites can insert local data, change photos and include contact information. For additional information regarding the use of these materials please contact your CDC epidemiologist.

## Appendix 3 – Test Used by STARHS Laboratory for Recency Classification

### BED HIV-1 Capture Enzyme Immunoassay

Currently, within STARHS the EIA is followed by the BED HIV-1 Capture Enzyme Immunoassay (BED; the assay that serves as the serologic marker of recent infection) which measures the concentration of anti-HIV IgG to total IgG in a sample. The result is reported as a standard optical density (SOD), a continuous measure that describes the relative concentration of anti-HIV IgG. If the SOD for a given sample is over a threshold predetermined to define a "long-standing" infection then the infection is deemed no longer recent. The period of time during which the SOD is below this threshold is termed the mean recency period. Although the mean recency period may vary slightly by HIV subtype, the mean recency period for calculating population-based incidence estimates in the United States is 162 days using the BED6.

The BED for recency testing is performed only on HIV antibody positive serum2,6 and is not approved as a diagnostic test. Because of the variability in antibody development among individuals, the predictive value of an individual's STARHS result is low; the results are reliable only as part of the population-based HIV incidence estimate. The Food and Drug Administration (FDA) has allowed the BED to be labeled "For Surveillance use only. Not for diagnostic or clinical use," therefore, STARHS results performed for purposes of HIS cannot be returned to individuals or their health care providers or used for clinical management. For more information on the BED please access the following link:

http://www.cdc.gov/hiv/topics/surveillance/resources/factsheets/pdf/bed.pdf

# Appendix 4.1

# STARHS Specimen Guidance: Specimens Originating from the Public Health Laboratory

HIV Incidence and Case Surveillance Branch
Division of HIV/AIDS Prevention
National Center for HIV, STD, Viral Hepatitis and TB Prevention
Centers for Disease Control and Prevention

Cleared by HICSB Management 10/15/2012

### *Transportation Overview*

In the transportation model where specimens originate from the public health laboratory (PHL), supplemental/confirmatory testing will have been performed at the state/local PHL. The state/local PHL will store all positive specimens received until STARHS testing eligibility is determined. All eligible specimens will be pulled, aliquoted into the designated cryogenic vials provided by the STARHS laboratory, relabeled with a STARHS identification number, and shipped to the STARHS laboratory by overnight shipping in accordance with the procedures described in Appendix 4.4 (*Guidelines for Preparing Specimens for Shipping and Transporting to the STARHS Laboratory*).

In this model, the PHL will submit laboratory report information to the state/local HIV surveillance system according to state HIV reporting requirements, but must also include the laboratory-assigned specimen accession number and the specimen collection date on the report.

**Figure 1** graphically depicts the flow of specimens and reports when specimens originate at the PHL and are then shipped to the STARHS laboratory.

### *Role of Public Health Laboratories*

The PHL must store all remnant HIV positive specimens until eligibility is determined by the HIV incidence surveillance staff from the state/local health department. The HIV incidence surveillance staff will provide the PHL with a list of all STARHS-eligible specimens, listed by specimen accession number. The PHL will then pull all eligible specimens and aliquot them into the designated cryogenic vial provided by the STARHS laboratory. The PHL will simultaneously re-label the eligible specimens with a STARHS identification number (ID) using labels provided to the PHL by the STARHS laboratory. The PHL must also provide the HIV incidence surveillance staff with a link between the STARHS identification number and the original specimen accession number. The PHL will ship all eligible specimens, labeled only with the STARHS identification number, to the STARHS laboratory according to the procedures described in the Appendix 4.4 (*Guidelines for Preparing Specimens for Shipping and Transporting to the STARHS Laboratory*).

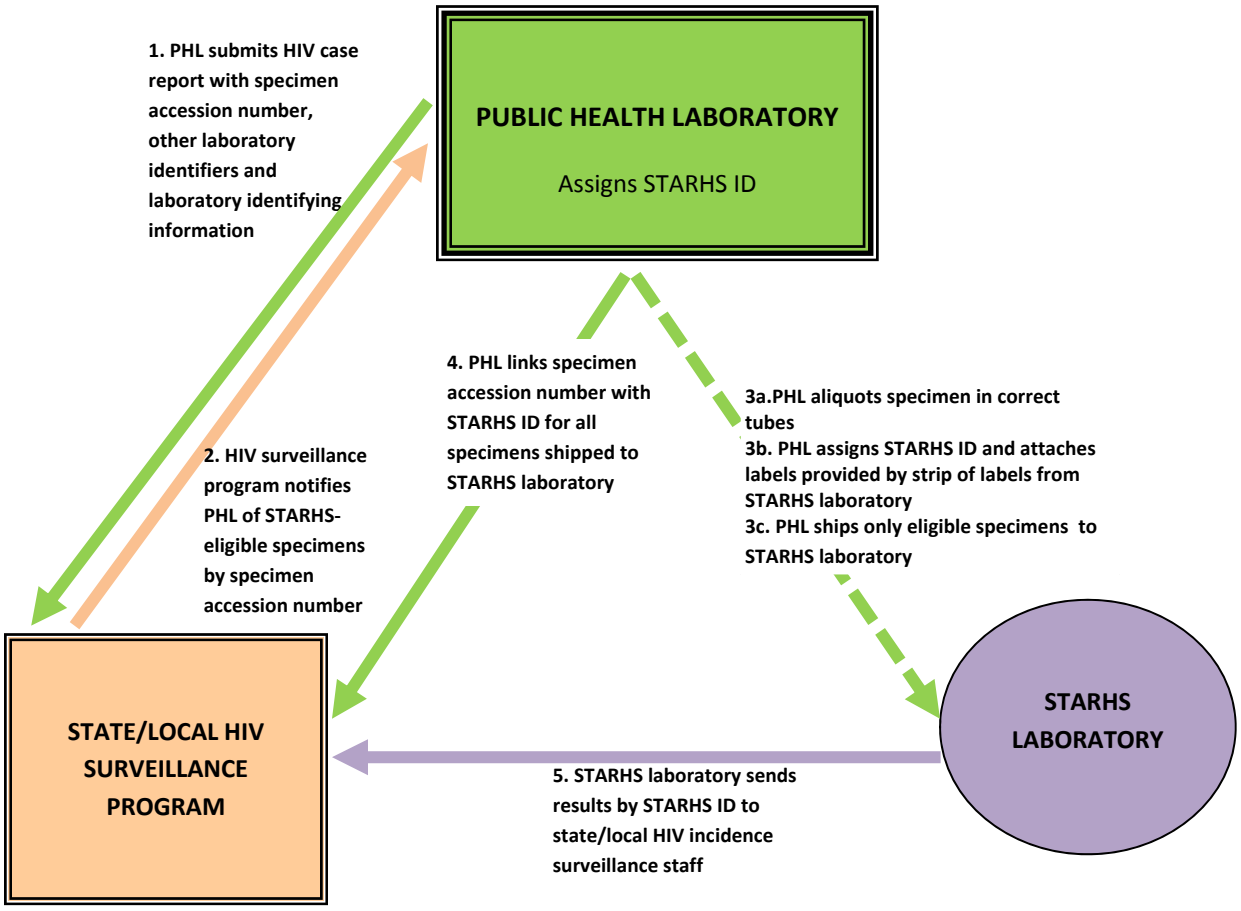
### *Role of State/Local HIV Incidence Surveillance Staff*

The state/local HIV incidence surveillance staff from the jurisdiction where patient specimens originated will determine STARHS eligibility and coordinate with the PHL to ensure that only eligible specimens are shipped to the STARHS laboratory. The state/local HIV incidence surveillance staff will also maintain the link, sent by the PHL, between the original specimen accession number and the STARHS number, and will manage STARHS results.

### *Role of STARHS Laboratory*

The STARHS laboratory will test all specimens received from a PHL (since only eligible specimens are sent) using the STARHS identification number. Once testing is complete, the STARHS laboratory will return results to the appropriate state/local health department.

**Figure 1**




Legend

Dotted line represents flow of specimen transport

Solid lines represent flow of information (e.g. laboratory information, STARHS laboratory, etc.)

Cleared by HICSB Management 10/15/2012

# Appendix 4.2

# STARHS Specimen Guidance: Specimens Originating from a Private/Commercial Laboratory and Sent to a Public Health Laboratory

HIV Incidence and Case Surveillance Branch
Division of HIV/AIDS Prevention
National Center for HIV, STD, Viral Hepatitis and TB Prevention
Centers for Disease Control and Prevention

Cleared by HICSB Management 10/15/2012

### *Transportation Overview*

In the transportation model where specimens originate from a private/commercial laboratory and are sent to the public health laboratory (PHL), confirmatory/supplemental testing will have been performed at a private/commercial laboratory. The private/commercial laboratory will send all HIV positive diagnostic specimens to the state/local PHL. The state/local PHL will store all specimens received from the private/commercial laboratories until STARHS eligibility is determined. All eligible specimens will be pulled, aliquoted into the designated cryogenic vials provided by the STARHS laboratory, relabeled with a STARHS identification number, and shipped to the STARHS laboratory by overnight shipping in accordance with the procedures described in Appendix 4.4 (*Guidelines for Preparing Specimens for Shipping and Transporting to the STARHS Laboratory*).

In this model, the originating laboratory will submit laboratory report information to the HIV surveillance system in accordance with state HIV reporting regulations, but must also include the laboratory-assigned specimen accession number, the specimen collection date, and testing laboratory identification on the report.

Many laboratories send enzyme immunoassay (EIA) positive specimens to a reference laboratory for confirmatory/supplemental testing which usually results in different laboratory accession numbers. In this case, care must be taken to ensure that the appropriate specimen accession numbers are associated with the correct HIV surveillance case report. **Figure 1** graphically depicts the flow of specimens and reports when specimens originate at a private/commercial laboratory and are sent to the PHL for storage prior to shipment to the STARHS laboratory.

### *Role of Private/Commercial Laboratories*

Private/commercial laboratories are responsible for forwarding two items for recency testing: (1) a laboratory report to the public health surveillance department per state HIV reporting regulations with specimen identifiers, specimen collection date, the laboratory-assigned specimen accession number, and testing laboratory identification information; (2) remnant HIV positive specimens to the public health laboratory from specimens labeled with the laboratory-assigned specimen accession number and testing laboratory identification information.

Private/commercial laboratories may elect to aliquot 0.5 ml of the remnant specimen to send to the PHL so that any additional portion of the remnant specimen may be stored at their facility. Also private/commercial laboratories may send the entirety of the remnant specimen, without any further manipulation, to the PHL.

### *Role of Public Health Laboratories*

The PHL must store all remnant HIV positive specimens shipped from a private/commercial laboratory until eligibility has been determined by the HIV incidence surveillance staff. HIV incidence surveillance staff will provide the PHL with a list of all STARHS-eligible specimens, listed by specimen accession number. The PHL will then pull all eligible specimens and aliquot them into the designated cryogenic vial provided by the STARHS laboratory. The PHL will simultaneously re-label the eligible specimens with a STARHS identification number (STARHS ID) using labels provided to the PHL by the STARHS laboratory. The PHL must also provide the HIV incidence surveillance staff with a link between the STARHS ID and the original

specimen accession number. The PHL will ship all eligible specimens, labeled only with the STARHS ID, to the STARHS laboratory according to the procedures described in the Appendix 4.4 (*Guidelines for Preparing Specimens for Shipping and Transporting to the STARHS Laboratory*).

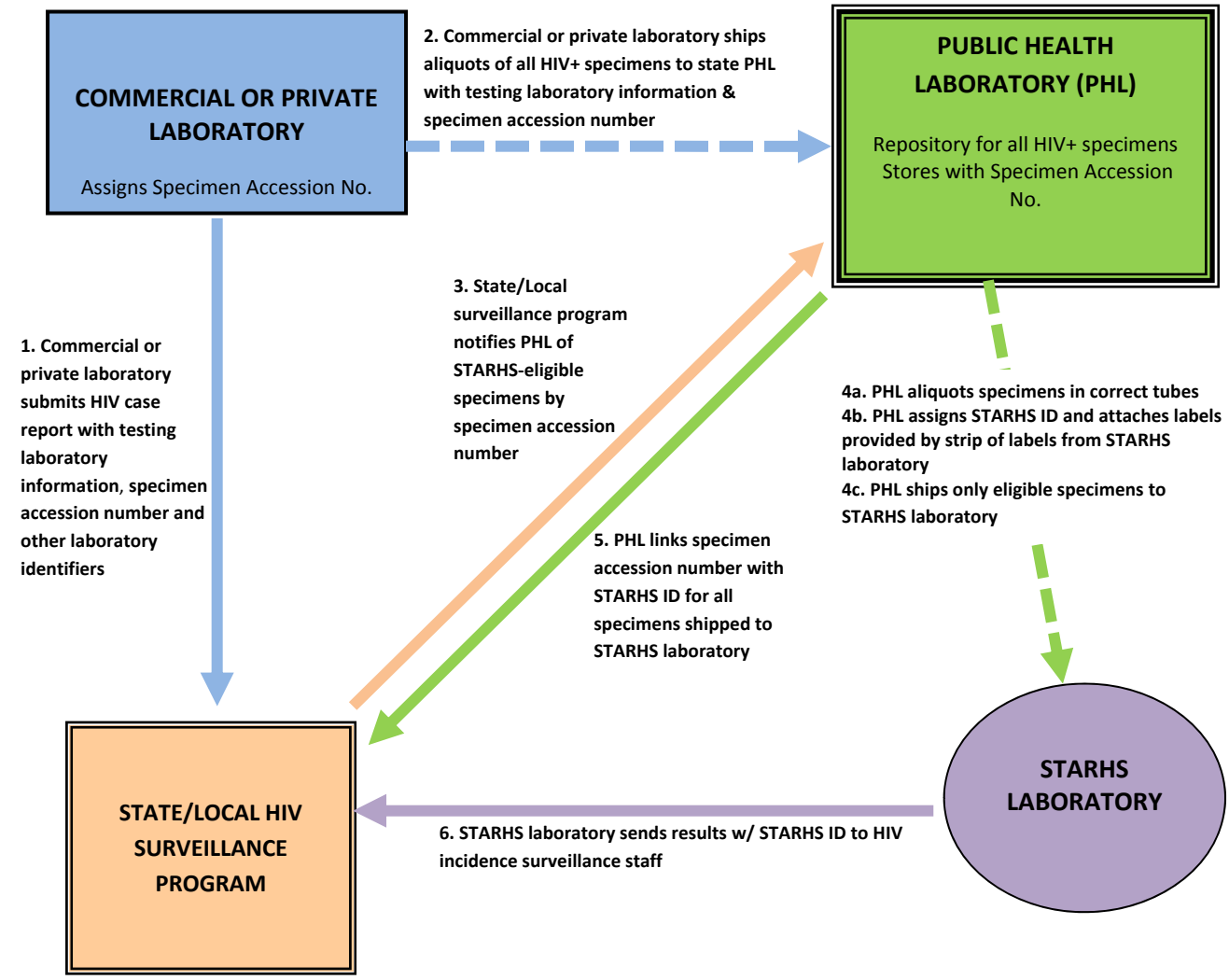
### *Role of State/Local HIV Incidence Surveillance Staff*

The state/local HIV incidence surveillance staff from the jurisdiction where patient specimens originate will determine STARHS eligibility and coordinate with the PHL to ensure that only eligible specimens are shipped to the STARHS laboratory. The state/local HIV incidence surveillance staff will also maintain the link between the original specimen accession number and the STARHS ID, and will manage STARHS results.

### *Role of STARHS Laboratory*

The STARHS laboratory will test all specimens received from a PHL (since only eligible specimens are sent) using the STARHS identification number. Once testing is complete, the STARHS laboratory will return results to the appropriate jurisdiction's state/ local HIV incidence surveillance staff.

**Figure 1.** Specimen Originates at a Private/Commercial Laboratory (for example, a University Hospital Laboratory, Regional or Local Independent Commercial Laboratory) and Sends Specimen to State Public Health Laboratory (serves as a pass-through facility)




Legend

Dotted lines represent flow of specimen transport

Solid lines represent flow of information (e.g. laboratory information, STARHS laboratory, etc.)

Cleared by HICSB Management 10/15/2012

# Appendix 4.3
# STARHS Specimen Guidance: Specimens Originating from a Private/Commercial Laboratories and Sent Directly to the STARHS Laboratory

HIV Incidence and Case Surveillance Branch
Division of HIV/AIDS Prevention
National Center for HIV, STD, Viral Hepatitis and TB Prevention
Centers for Disease Control and Prevention

Cleared by HICSB Management 10/15/2012

### *Transportation Overview*

In the transportation model where specimens originate from a private/commercial laboratory and are sent directly to the STARHS laboratory, the originating private/commercial laboratory performing the supplemental/confirmatory testing will send all remnant sera/plasma from HIV positive diagnostic specimens directly to the STARHS laboratory (by-passing the state /local PHL) by overnight shipping in accordance with the procedures described in Appendix 4.4 (*Guidelines for Preparing Specimens for Shipping and Transporting to the STARHS Laboratory*).

The STARHS laboratory will store specimens until STARHS eligibility is determined by the state/local HIV incidence surveillance staff, at which point, the STARHS laboratory will pull the specimens, aliquot, re-label them with a STARHS identification number, and perform recency testing.

In this model, the originating laboratory will submit case report information to the public health surveillance department according to state HIV reporting requirements, but must also include the laboratory-assigned specimen accession number, the specimen collection date, and testing laboratory identification on the report.

Many laboratories send enzyme immunoassay (EIA) positive specimens to another reference laboratory for confirmatory/supplemental testing, which usually results in different laboratory accession numbers. In this case, care must be taken to ensure that the appropriate specimen accession numbers are associated with the correct HIV case surveillance report. **Figure 2** depicts the flow of specimens and reports when specimens originate at a private/commercial laboratory and are then shipped directly to the STARHS laboratory.

### *Locating Specimens*

Laboratories that perform routine diagnostic supplemental/confirmatory HIV testing or immunological status tests such as CD4, or viral load counts should report results to the state/local health department surveillance program per existing state HIV reporting requirements.

In each surveillance area, all laboratories that perform HIV diagnostic tests for residents of the surveillance area should be identified from a review of local HIV surveillance data and state laboratory licensing records. These laboratories should be contacted to request that remnants of HIV-positive diagnostic specimens be made available for recency testing. The reports that these laboratories send to the state/local health department surveillance programs per the established reporting requirements should include specimen accession numbers. Some laboratories perform only the initial screening HIV EIA test, and then send the specimen to a reference laboratory for confirmatory HIV testing. Originating laboratories are those where a specimen is first sent for testing. Reference laboratories are those where a specimen is sent for confirmatory testing when the originating laboratory does not do confirmatory testing. A specimen may be assigned a different accession number at each laboratory that receives it; therefore it is important to ensure that the accession number that is assigned by the laboratory that ships the specimen to the STARHS laboratory is transmitted to the surveillance program as part of the laboratory reporting

process. Surveillance areas should maintain a directory of laboratory contacts at all reporting laboratories to facilitate communication in the event that reporting or shipping of specimens is disrupted or that changes in policy or procedures need to be communicated.

### *Determining the Disposition of Specimens and Communicating with the Public Health and STARHS Laboratory*

Private/commercial laboratories that ship directly to the STARHS laboratory will send all HIV positive specimens. The area HIV incidence surveillance staff, using routine HIV surveillance reporting procedures, determines whether the specimen represents the person's first reported positive HIV test result in the HIS area and is eligible for recency testing.

The HIS program must inform the STARHS laboratory of the disposition of the specimen because in this transportation model a remnant sample of every HIV positive diagnostic specimen will be shipped to the STARHS laboratory.

### *Preparing the STARHS Eligibility Lists*

The STARHS laboratory receives shipments of HIV-positive specimens from private and commercial laboratories. Upon receipt, the specimens are checked against the shipping manifest to verify that the specimen accession numbers on the tubes match those listed on the manifest. The verified accession numbers are then entered into the STARHS specimen database by 1) manually entering individual accession numbers, 2) barcode scanning, if tube labels contain barcodes, or 3) file importing, if the manifest is provided as an Excel file. Once the specimens have been accessioned into the database, a list is generated and imported into the STARHS freezer inventory system to link the freezer location of each specimen to the accession number in the database.

HIV incidence surveillance staff for each site determines the eligibility of the specimens and provides this information to the STARHS laboratory in the form of an eligibility list. Upon receipt of an eligibility list, the STARHS laboratory uploads the list into the specimen database and searches for the specimen accession numbers that match those on the eligibility list. As matches are identified, the eligible specimens are located and retrieved from the freezer, and recency testing is conducted.

In order for the matching process to be successful, eligibility lists that are sent by incidence coordinators to the STARHS laboratory must be produced using a standard format (Table 1). A standard template for preparing eligibility lists has been developed as an Excel file named '**2011 CDC STARHS Eligibility list_v0511**'. The first tab of this Excel spreadsheet is the template for producing eligibility lists. The second tab contains the list of standard codes for the state/site abbreviations and laboratory names. The organization of the spreadsheet and the contents must adhere to this standard format to allow for uploading of the data into the STARHS specimen database at the STARHS laboratory.

**Table 1**. STARHS Eligibility List Template

| Accession # | 2nd Accession # Format (If applicable) | **Lab Name** (Use standard lab names only) | **State** (Use standard codes only) | **Eligibility** Record 'Eligible' or 'Ineligible' (Sort by eligibility status) | **Previously Requested?** (Indicate only if answer is 'Yes') |
|---|---|---|---|---|---|
| | | | | Eligible | |
| | | | | Ineligible | |
| | | | | | |
| | | | | | |
| | | | | | |

HIV incidence surveillance staff should use only the template from the **2011 CDC STARHS Eligibility list_v0511.xls** provided by the STARHS laboratory for producing eligibility lists. The template should not be modified. The order of the columns must remain as it appears in the template or the file will not function properly during the upload process.

For each specimen accession number listed in the **Accession Number** column, the **State** and **Eligibility** columns must be completed. There should be no empty cells in these columns. Columns entitled **2nd Accession Number**, **Lab Name,** and **Previously Requested** should be completed if applicable, but may have empty cells. See details for each column below.

**State:** In the **State** column, HIV incidence surveillance staff should use only the standard code for their site listed in the 'standard codes' tab of the spreadsheet (Table 2). This code should be completed for each accession number on the list.
**Eligibility:** In the **Eligibility** column, HIV incidence surveillance staff should list either 'Eligible' or 'Ineligible' for each accession number on the list. The terms 'Test' and 'Toss' have been used by some in the past, but these should no longer be used.
**Lab Name:** In the **Lab Name** column, HIV incidence surveillance staff should use only the abbreviated laboratory name that is listed in the 'standard codes' tab of the spreadsheet (Table 3). The laboratory name should only be added if it appears on the list. If a new laboratory comes on board and an abbreviated laboratory name has not been added to the list yet, the **Lab Name** cell should be left blank for that accession number. The list will be updated, as needed, and notification of updated laboratory names will be provided.
**2nd Accession Number:** If there is a second accession number associated with a particular specimen, that number should be added to the **2nd Accession Number** column. If there is no additional accession number for a specimen, this cell may be blank.
**Previously Requested:** In general, once a specimen accession number has appeared as 'Eligible' or 'Ineligible' on an eligibility list, there should be no need to include it on future lists. However, if results have not been received after 3 months, it may be necessary to request the specimen again using the eligibility list. If an accession number that was requested on a previous eligibility list is included, HIV incidence surveillance staff should indicate 'Yes' in the

**Previously Requested** column. In some cases the format of the accession number in the eligibility list may not match exactly with the number in the STARHS laboratory specimen database. By indicating 'Yes' in the **Previously Requested** column, additional procedures will be implemented to try to locate the specimen for testing.

### *Organizing the Eligibility List File for Sending to the STARHS laboratory*

After the eligibility list has been completed, the entire list should be sorted using the **Eligibility** column so that all of the 'Eligible' specimens are grouped together and all of the 'Ineligible' specimens are grouped together.

The file should be encrypted as a PGP zip file, preferably as a self-decrypting PGP zip file to allow the file to be opened on a computer that doesn't have PGP software. The STARHS laboratory does not use the PGP key system.

Each site should use their site's standard password to protect their files. This makes it much easier if the STARHS laboratory needs to re-open the initial list on the CD at a later time.

The encrypted file must be saved to a CD, and labeled with the site name and the date. The STARHS laboratory receives disks from multiple sites and it is important for the disk to be labeled so that it can be easily identified.

### *Untested Specimen Lists*

An **Untested Specimen List** is produced and distributed by the STARHS laboratory to all sites at least quarterly. This list contains the accession numbers of specimens that have been received by the STARHS laboratory and have not been tested yet.

HIV incidence surveillance staff may request testing of a specimen listed on the **Untested Specimen List**, using the **Eligibility List** template following all standard procedures for preparing the eligibility list**,** as described above.

The specimen accession number is recorded in the **Eligibility List** template exactly as it appears on the **Untested Specimen List**. The list of accession numbers on the **Untested Specimen List** is generated directly from the STARHS laboratory specimen database. This will ensure that the accession number on the eligibility list is in the same format as it appears in the database, allowing it to be readily located and tested.

**Table 2.** List of Standard Codes for State/Site Abbreviations

| Code | State/Site | | Code | State/Site |
|---|---|---|---|---|
| AZ | Arizona DOH | | WA | Washington DOH |
| CHI | Chicago DOH | | MI | Michigan DOH |
| CT | Connecticut DOH | | PHI | Philadelphia DOH |
| FL | Florida DOH | | SC | South Carolina DOH |

| | | | | |
|---|---|---|---|---|
| HOU | Houston DOH | | LAC | L A County DOH |
| LA | Louisiana DOH | | SAN | San Francisco DOH |
| MA | Massachusetts DOH | | DC | Washington DC DOH |
| MS | Mississippi DOH | | NC | North Carolina DOH |
| NJ | New Jersey DOH | | SEA | Seattle DOH |
| NYC | NY City DOH | | CO | Colorado DOH |
| NYS | NY State DOH | | IN | Indiana DOH |
| TX | Texas DOH | | CA | California DOH |
| VA | Virginia DOH | | AL | Alabama DOH |
| | | | | |

**Table 3.** List of Standard Codes for Laboratory Names

| No. | Standard Name | Laboratory |
|---|---|---|
| 1 | ARUP | ARUP Laboratories |
| 2 | MAYO | Mayo Medical Laboratories |
| 3 | QUEST | Quest Diagnostics |
| 4 | BIO-REFERENCE | Bio Reference Laboratory |
| 5 | PRINCETON | Princeton Biomedical Laboratories |
| 6 | SPECIALTY | Specialty Laboratories, Inc. |
| 7 | LABCORP | Laboratory Corporation of America |
| 8 | WARDE | Warde Medical Laboratory |
| 9 | SUNRISE | Sunrise Medical Laboratory |
| 10 | NASSAU | Nassau County Dept of Health Lab |
| 11 | UNC | University of North Carolina @ Chapel Hill |
| 12 | UVAHS | University of Virginia Health System |
| 13 | CRL | Clinical Reference Laboratory |
| 14 | HERITAGE | Heritage Labs International |
| 15 | MCV | Virginia Commonwealth University Health System |
| 16 | DUKE | Duke University Medical Center ,Durham NC |
| 17 | ECMC | Erie County Medical Center |
| 18 | SBUMC | Stony Brook University Medical Center |
| 19 | KALEIDA | Kaleida Health |
| 20 | STLUKES | St Luke's Roosevelt Laboratory |
| 21 | WESTCHESTER | Westchester County Labs & Research |
| 22 | MARICOPA | Maricopa Integrated Health System, Phoenix AZ |
| 23 | BSI | Bendiner & Schlesinger Inc. |
| 24 | PAML | Pathology Associates Medical Laboratories |

| | | |
|---|---|---|
| 25 | UNYTS | Upstate New York Transplant Services |
| 26 | ARC | National Confirmatory American Red Cross |
| 27 | KAISER | Kaiser |
| 28 | SFGH | San Francisco General Hospital |
| 29 | CDD | Center for Disease Detection, San Antonio TX |
| 30 | ALAMEDA | ALAMEDA County Public Health Lab, Oakland CA |
| 31 | SANDIEGO | San Diego County Public Health Laboratory, CA |

## Role of Private/Commercial Laboratories

The private/commercial laboratories are responsible for forwarding two items for recency testing: (1) a laboratory report to the HIV surveillance system per local requirements, including the collection date, the laboratory-assigned specimen accession number, and identification information about the testing facility; (2) remnant HIV positive serum/plasma from diagnostic specimens labeled with the laboratory-assigned specimen accession number.

The private/commercial laboratories may elect to aliquot 0.5 ml of the remnant specimen for shipment to the STARHS laboratory so that any additional portion of the remnant specimens may be stored at their facility, or they may send the entirety of their remnant specimens, without any further manipulation, to the STARHS laboratory.

## Role of STARHS Laboratory

The STARHS laboratory must store all remnant HIV positive specimens received until eligibility has been determined by the appropriate jurisdiction's HIV Incidence Surveillance staff. The HIV incidence surveillance staff will provide the STARHS laboratory with a list of all specimens eligible to be tested for recent infection, listed by specimen accession number. If eligibility has not been determined after 2 years from the time the specimen was received in the STARHS laboratory, the specimens may be destroyed according to established laboratory methods.

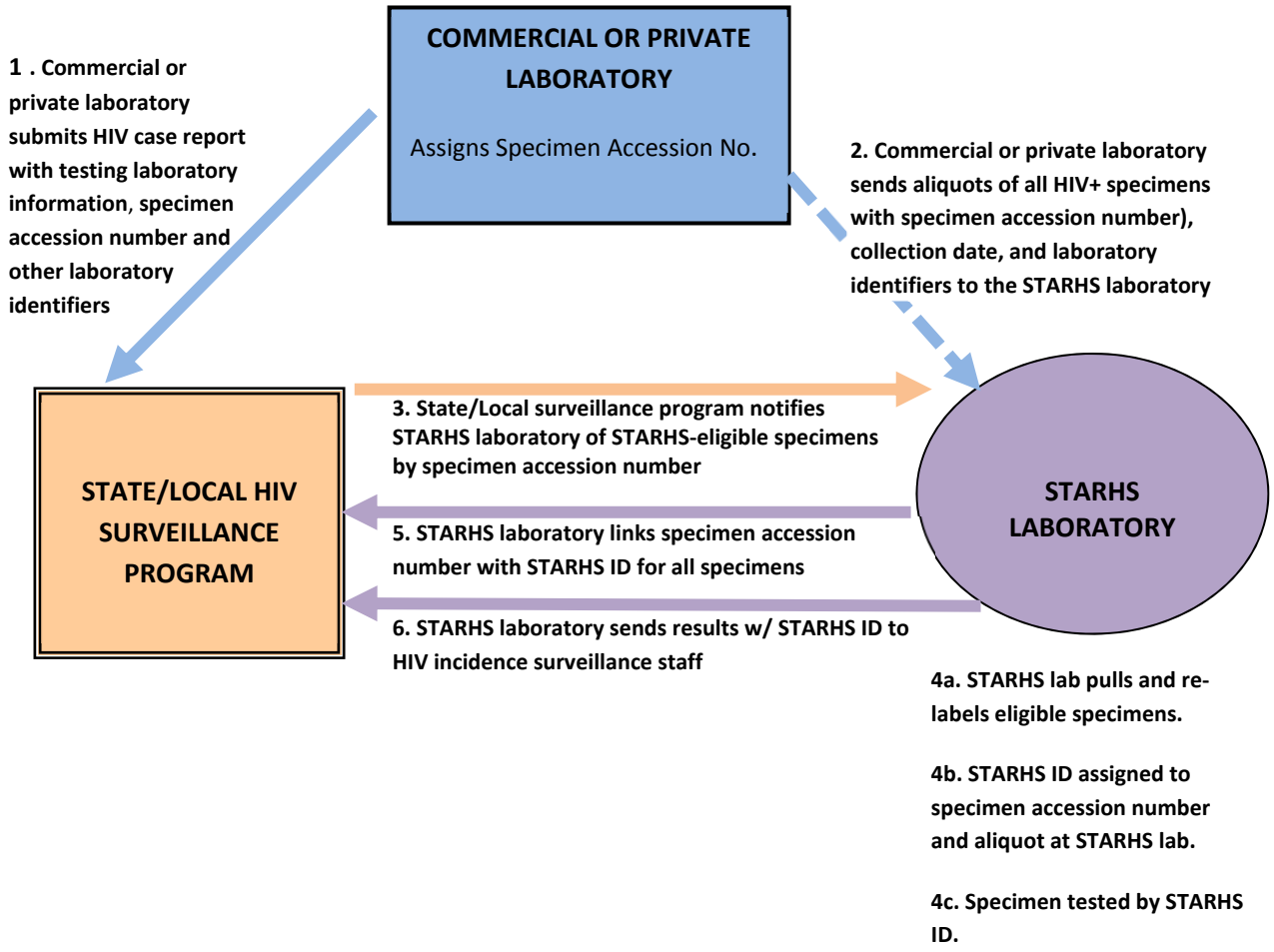
The STARHS laboratory will pull all eligible specimens and aliquot them into the designated cryogenic vial for testing. The STARHS laboratory will simultaneously re-label the eligible specimens with a STARHS identification number (STARHS ID). The STARHS laboratory will provide the appropriate HIV incidence surveillance staff with a link between the STARHS identification number and the original specimen accession number. All links at the STARHS laboratory are maintained on a secure internal network that ensures that confidentiality is protected and maintained to meet the standards for HIV surveillance listed in National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention's (NCHHSTP) ***Data Security and Confidentiality Guidelines for HIV, Viral Hepatitis, Sexually Transmitted Disease, and Tuberculosis Programs: Standards to Facilitate Sharing and Use of Surveillance Data for Public Health Action.*** The STARHS laboratory will test all eligible specimens and send results (using only the STARHS ID) to the HIV incidence surveillance staff from the appropriate jurisdiction. The STARHS results are for surveillance purposes only therefore results will not be reported back to the originating laboratory, provider, or client.

The STARHS laboratory will send each HIV incidence surveillance jurisdiction an "Untested Specimens List" at least quarterly. HIV incidence surveillance staff should examine the list to identify remnant HIV-positive specimens sent directly to the STARHS laboratory and representing diagnostic specimens from newly reported cases of HIV infection. These specimens should be listed on the STARHS Eligibility List.

## Role of the State/Local HIV Incidence Surveillance Staff

State/local HIV incidence surveillance staff from the jurisdiction where a specimen originated will determine STARHS eligibility and coordinate with the STARHS laboratory to ensure that proper specimens are pulled and tested. State/local HIV incidence surveillance staff will also maintain the link between the original specimen accession number and the STARHS ID number, and will manage the STARHS results.

**Figure 1.** Specimen Originates at National Commercial Laboratory or Smaller Private Laboratory and is Sent Directly to the STARHS Laboratory




Legend

Dotted line represents flow of specimen transport

Solid lines represent flow of information (e.g. laboratory information, STARHS laboratory, etc.)

Cleared by HICSB Management 10/15/2012

# Appendix 4.4

# Guidelines for Preparing Specimens for Shipping and Transporting to the STARHS Laboratory

HIV Incidence and Case Surveillance Branch

Division of HIV/AIDS Prevention

National Center for HIV, STD, Viral Hepatitis and TB Prevention

Centers for Disease Control and Prevention

Cleared by HICSB Management 10/15/2012

# Guidelines for Preparing Specimens for Shipping and Transporting to the STARHS Laboratory

## *Purpose*

This standard operating procedure describes methods for the handling, storage, and shipping of serum/plasma specimens that will be tested for recent HIV-1 infection using STARHS. Results from these tests will help estimate HIV incidence.

## *Introduction*

Surplus serum/plasma from positive HIV diagnostic specimens is to be collected, frozen, and shipped to the STARHS laboratory. Public health laboratories (PHL) that ship only specimens eligible to be tested for recent infection must use vials and labels specified and supplied by the STARHS laboratory. Commercial/private laboratories that ship all remnant HIV-positive specimens to the STARHS laboratory may choose to freeze and ship specimens using the original primary specimen tube or aliquot tubes provided by the commercial/private laboratory and labeled with that laboratory's specimen accessioning label.

## *Type of Specimens Shipped to STARHS Laboratory*

HIV positive serum or plasma from confirmed HIV positive diagnostic samples will ultimately be shipped to the STARHS laboratory.

## *Specimen Volume*

The optimal quantity of serum/plasma required for recency testing is 0.5ml per aliquot. However, if less than 0.5ml of the remnant sample is available for recency testing the sample should still be sent to the STARHS laboratory. **The STARHS laboratory is the only laboratory that should determine if a sample should be rejected due to insufficient quantity**.

## *Sample Storage*

Short-term (less than one week) storage of samples in the refrigerator (2° – 8°C) is acceptable, but for long-term storage (more than one week), samples must be frozen at -20°C or colder. Long-term storage includes any period of time that the samples are kept at the originating/testing laboratory or a "pass through" PHL prior to shipment to the STARHS laboratory, or the interim period while STARHS eligibility is being determined. Effort should be made to avoid repeated freezing and thawing of samples, as repeated freeze-thaw cycles may lead to erroneous results.

The freezer must contain adequate space to store specimens. It is recommended that, if not already in practice, a daily temperature log be kept to ensure the freezer is operating properly. The freezer should be housed in a location with proper ventilation to avoid overheating and freezer failure.

### *Specimen Numbering*

The specimen number on the samples shipped to the STARHS laboratory will either be the original laboratory-assigned specimen accession number or the STARHS identification number (STARHS ID) depending on the transport model. Detailed information about specimen numbering is included in Appendices 4.1-4.3 of this guide.

### *Specimen Retention*

State/local HIV incidence surveillance staff must coordinate with the laboratory storing HIV positive remnant sera (the STARHS laboratory and/or the state/local PHL) to identify eligible samples to be tested for recent infection using STARHS. However, not all stored samples will be eligible for recency testing, and will therefore have to be identified for disposal. State/local HIV incidence surveillance staff and the storing laboratory should communicate regularly (every 1-3 months) to discuss any specimens that were not on the eligibility list to determine if the sample can be disposed of or if further investigation is needed. Samples should not be destroyed or disposed of until eligibility is definitively determined. Specimens at the STARHS laboratory that have not been requested for testing after a period of two years, may be discarded.

### *Shipping Program*

Since 2008, the STARHS laboratory has directed the shipping and transport of specimens through the STARHS specimen shipping program. The STARHS specimen shipping program includes the provision of shipping materials and a courier account for overnight shipping of specimens to the STARHS laboratory. All laboratories, public health, commercial and private, that ship specimens to the STARHS laboratory are eligible to utilize this shipping program. Specimens must be shipped on dry ice by same-day or overnight shipping. Specimens may be shipped from originating labs to the state PHL which may serve as a "pass through" facility or may be shipped to the STARHS laboratory as 'Diagnostic Specimens.' All laboratories shipping HIV positive samples must be certified to ship dangerous goods and observe the dry ice requirement, regardless of where specimens are shipped.

The STARHS laboratory will provide shipping materials (shipping containers, bags/containers and absorbent materials required per federal regulations, as well as fiberboard specimen boxes and all required shipping labels) and a courier account number for overnight shipping of specimens. Dry ice will not be provided by the STARHS laboratory. Each participating laboratory must identify at least one contact person for shipping samples. The STARHS laboratory will provide instructions to this individual on how to access the courier account and utilize the system.

### *Frequency of Shipments*

The frequency of specimen shipments to the STARHS laboratory will be determined by the shipping laboratory. To determine this frequency the shipping laboratory should consider factors such as specimen retention policies and freezer/storage space and consult with the STARHS laboratory delivery service to ensure that specimens do not thaw in transit.

### *Additional Information for All Other Private Laboratories*

Intra-state shipments from private laboratories to the PHL may be shipped by Federal Express (or a similar commercial courier) or an established local courier service. Funding permitting, sites may elect to set-up a billing account with Federal Express (or a similar commercial courier) to off-set costs incurred by the private laboratory for shipping to the state/local PHL.

### *Tracking Shipments*

The shipping laboratory should notify the receiving laboratory (state PHL or STARHS laboratory) by fax or email when specimens are shipped, including the tracking number. The receiving laboratory will be responsible for tracking the shipments and will notify the originating laboratory if the specimens are not received.

### *Sample Rejection Criteria*

Sample rejection due to thawing, breakage, insufficient quantity, or lost-in-transit will be determined and recorded by the STARHS laboratory and will be reported to the appropriate HIV surveillance system.

### *Confidentiality and HIPAA Regulations*

STARHS testing must ensure that confidentiality is protected and maintained to meet standards for HIV surveillance. The Privacy Rule of the Health Insurance Portability and Accountability Act (HIPAA) regulations permit protected health information to be shared for the purposes of public health surveillance activities, allowing the originating laboratories to send specimens labeled with their laboratory assigned accession number to either the state PHL or the STARHS laboratory, where eligible samples will be re-assigned a unique STARHS identification number prior to testing for recent infection. This process will minimize re-labeling errors and also simplify the shipment procedures for private laboratories. However, the state/local public health department must have the laboratory accession number to link the test result to the patient information in the surveillance record. Therefore, the laboratory accession number must be included on the HIV laboratory report form sent to the state/local public health department by the originating laboratory.

### *Preparing and Packing Specimens for Shipping to the STARHS Laboratory*

**Setting and personnel for specimen processing**

1. Centrifugation, aliquoting, and shipping should be performed at or under the auspices of a laboratory that is CLIA- certified for handling HIV-positive specimens.
2. All personnel handling specimens should receive blood borne pathogens training. See **OSHA's Occupational Exposure to Bloodborne Pathogens Standard:** http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10051.

3. Personnel handling or processing specimens should have appropriate laboratory training in the relevant laboratory techniques for handling HIV-positive specimens.
4. The setting in which centrifugation, aliquoting, and shipping occurs should meet Biosafety level 2 specifications required by the U.S. Department of Health and Human Services for handling of specimens containing HIV [http://www.cdc.gov/od/ohs/biosfty/bmbl4/bmbl4toc.htm, Biosafety in Microbiological and Biomedical Laboratories (BMBL) 5th Edition].

**Required Materials**

1. Cryogenic vials - Supplied by STARHS laboratory to PHL only.
2. Specimen labels - Supplied by STARHS laboratory to PHL only. Label will identify the sample (includes barcode and number).
3. Cardboard storage boxes for cryogenic vials with divider insert – Supplied by STARHS laboratory to all laboratories except when not wanted.
4. A supply of dry ice in pellet form - not supplied by STARHS laboratory.
5. Insulated shipping containers certified to ship frozen diagnostic specimens (HIV+ serum and dry ice) - Supplied by STARHS laboratory.
6. STARHS/FedEx account access - Supplied by STARHS laboratory.
7. Materials for shipper packing - Supplied by STARHS laboratory.
8. STARHS specimen submission form.

**Specimen Collection and Processing**

1. All processing of specimens should be done by personnel qualified to handle HIV-positive specimens under the auspices of a laboratory equipped for the handling of HIV-positive specimens [see http://www.cdc.gov/od/ohs/biosfty/bmbl4/bmbl4toc.htm, Biosafety in Microbiological and Biomedical Laboratories (BMBL) 5th Edition ].

1.1 Aliquot the serum (0.5 ml per cryogenic vial). Use labels to identify the specimen and record this information in the proper setting. (Specimen log for eventual transfer to access laboratory database for linked samples and temporary database for unlinked specimens)

1.2 Store aliquots in refrigerator or freezer until eligibility requirements are met and scheduled shipping date has arrived. Samples for recency testing can be refrigerated at 2-8°C, but for long term storage and shipping, samples should be frozen at -20°C.

1.2.1 It is recommended that, if not already in practice, a daily temperature log be kept to ensure the freezer is operating properly.

1.2.2 The freezer should be housed in a location with proper ventilation to avoid overheating and freezer failure.

1.2.3 Staff must ensure adequate freezer space to store specimens.

**Shipping**

1. Specimens for recency testing should be sent to the STARHS laboratory using the address provided on the STARHS specimen submission form and under STARHS laboratory in this document. All specimens will be shipped as diagnostic specimens using International Air Transport Association (IATA) Packing Instructions 650. Dry ice will be included with each shipment using IATA Packing Instructions 904.

   1.1 Because samples will be shipped using dry ice, shipping personnel must be trained and certified to ship dangerous goods. See below for a list of companies that provide training.

   1.2 Establish contact with N'ko Lea Ali-Napo (nla01@health.state.ny.us) at the STARHS laboratory.

   1.3 An invitation to set up a STARHS / FedEx account for the laboratory will be provided by the STARHS laboratory.

   1.4 A fax or e-mail system should be set up to notify the STARHS laboratory of incoming shipment.

   1.5 Adequate STP320 or equivalent shipping containers must be available. The STARHS laboratory will return them to the shipping laboratory. These are expensive shippers and need to be re-used.

   1.6 Laboratories should ensure access to the STARHS FedEx account and that they are able to print shipping courier air bills.

2. Packing and shipping dates should be determined in consultation with N'ko Lea Ali-Napo at Wadsworth Center Bloodborne Viruses Laboratory. It is recommended that staff at the laboratory read and walk through all of the following steps **before** starting the preparation of the actual shipment to ensure familiarity with the shipping procedures and requirements. Specimens should not be shipped on Friday or on the day before an official holiday. **Specimens should remain frozen at all times and therefore should not arrive at the Wadsworth Center when the center is closed.**

### *STARHS Medium Shipper Instructions*

The medium shipper will accommodate 2 fiberboard freezer boxes that will hold up to 81 specimens in each box for a maximum total of 162 specimens. All unused freezer boxes, plastic bags, and Tyvek® envelopes should be left in the box when shipping specimens to the STARHS laboratory.

Supplies included in each shipper:

1 - Category B label

1 - UN1845 dry ice label

1 - Class 9 hazard miscellaneous goods label

2 - Fiberboard freezer boxes with divider insert

2 - Large Inner leak proof polybags with absorbent strip (instructions printed on polybag)

2 - Large Tyvek® outer envelopes (instructions printed on envelope)

## DIRECTIONS TO PACK MEDIUM SHIPPER

1. Load specimen tubes into the box in the order listed on the manifest.
2. Place the box in the polybag and seal it according to the instructions on the polybag.
3. Insert the polybag and box into the Tyvek® envelope and close the envelope according to the instructions on the envelope
4. Return both fiberboard specimen boxes regardless of whether they have been used.
5. If both boxes are not used, place the empty box on the bottom of the insulated box and place the box containing specimens on top.
6. Add the appropriate amount of dry ice.
7. Include a copy of the STARHS Specimen submission form and manifest containing the specimen list.
8. Seal the closed box with shipping tape
9. Place labels (Category B, UN1845, and Class 9) and the FedEx airbill on the box

## STARHS Large Shipper Instructions

The large shipper will accommodate 4 fiberboard freezer boxes that will hold up to 81 specimens in each box for a maximum total of 324 specimens. Please leave all unused freezer boxes, plastic bags, and Tyvek® envelopes in the box when shipping specimens to the STARHS laboratory.

Supplies included in each shipper:

1 - Category B label

1 - UN1845 dry ice label

1 - Class 9 hazard miscellaneous goods label

4 - Fiberboard freezer boxes with divider insert

4 - Large Inner leak proof polybags with absorbent strip (instructions printed on polybag)

4 - Large Tyvek® outer envelopes (instructions printed on envelope)

## DIRECTIONS TO PACK LARGE SHIPPER

1. Load specimen tubes into the box in the order listed on the manifest.
2. Place the box in the polybag and seal the polybag according to the instructions on the polybag.
3. Insert the polybag and box into the Tyvek® envelope and close the envelope according to the instructions on the envelope.
4. Stack the envelopes with the boxes of specimens into the shipping box.
5. Add the appropriate amount of dry ice around the boxes.
6. Place the foam plug over the specimens.
7. Place any unused supplies on top of the foam.
8. Include a copy of the STARHS Specimen submission form and manifest containing the specimen list.
9. Seal the closed box with shipping tape.
10. Place labels (Category B, UN1845, and Class 9) and the FedEx airbill on the box

NOTE: In order to conserve supplies it is requested that unused supplies be returned with each shipment.

Questions should be directed to **Thomas T. Miller**, ttm01@health.state.ny.us at **(518) 474-2163.**

**Important STARHS shipping information**

**Shipping address**

For all specimen shipments for recency testing please be sure to use the correct shipping address. This correct address is printed on the STARHS shipping manifest. Using other addresses could lead to delays or lost shipments. Specimen shipments should be sent to:

**NYSDOH Wadsworth Center**
Axelrod Institute
Bloodborne Viruses Laboratory: STARHS
120 New Scotland Avenue
Albany, NY 12208
Attn: N'ko Lea ALI-NAPO

**Email notifications**

Notifications of specimen shipments and transmission of manifests from commercial laboratories by email should be sent using the following email addresses only:

nla01@health.state.ny.us (Lea Ali-Napo)

wac01@health.state.ny.us (William Carmichael)

bgranger@wadsworth.org (Brian Granger)

Both of these email addresses should be included when sending an email notification of shipment. Do not include any other email addresses previously associated with the STARHS laboratory when sending shipping manifests. It is important that the manifests and other email notifications regarding shipments be directed only to current staff at the STARHS laboratory.

**FAX notifications**

FAX notifications of upcoming STARHS shipments should be addressed to **Lea Ali-Napo** or **William Carmichael**. Please continue to use the FAX number **(518) 473-0008** for this purpose.

**Primary contact for STARHS**

N'ko Lea Ali-Napo (Lea), nla01@health.state.ny.us

## STARHS Shipment Contact

STARHS / FedEx account access

Thomas T. Miller, ttm01@health.state.ny.us

Shipment notification and shipment

N'ko Lea Ali-Napo, nla01@health.state.ny.us

William Carmichael, wac01@health.state.ny.us

Brian Granger, bgranger@wadsworth.org

## Shipping Address

**NYSDOH Wadsworth Center**
Axelrod Institute
Bloodborne Viruses Laboratory: STARHS
120 New Scotland Avenue
Albany, NY 12208

Phone (518) 474-2163
FAX (518) 473-0008

## Director of Blood-borne Diseases Laboratory

**Monica M. Parker, Ph.D., mmp09@health.state.ny.us**

***STARHS Specimen Submission Form***

Please complete this form and send it with each shipment. Specimens should be sent to:

**NYSDOH Wadsworth Center**
**Axelrod Institute**
Bloodborne Viruses Laboratory: STARHS
120 New Scotland Avenue
Albany, NY 12208
Attn: N'ko Lea ALI-NAPO

SHIPPING FACILITY INFORMATION:

Name:_____ _____

Address:_____

_____

Phone Number:_____

Fax:_____

Email:_____

Contact Person:_____

RESULTS SENT TO:

Name_____

Address_____

_____

Phone Number:_____

Fax_____

Email:_____

☐ **INCIDENCE SURVEILLANCE (HICSB) - List of eligible specimens sent separately**

Please place the labels on the tubes with left side even with the bottom of the tube threads.

**RANGE OF SPECIMENS SENT:**

_____

_____

_____

_____

FOR NYS DOH USE:

| | |
|---|---|
| Date Rec'd. | Date Accessioned: |
| Date Tested: | Date Reported: |
| Date Specimens Returned: | Date Shipping Container(s) Returned: |

Cleared by HICSB Management 10/15/2012

**Training and Certification for Shipping Infectious Substances**

The following companies provide training for dangerous goods shipping, however, the Centers for Disease Control and Prevention does not endorse any particular company.

FedEx 800-GO-FEDEX 3 day IATA based training Covers all hazardous materials Cost is $550

Saf-T-Pak 800-814-7484 specifically for infectious and diagnostic substances, and dry ice 3 options---One day seminar, On-site programs, or Interactive CD For interactive CD: for one sitting, can be done in 3-5 hours. Certificate good for 2 years OR until regulations change Cost is ~$250

Viking Packaging (Oklahoma) 800-788-8525—David Weilert Seminars monthly in Tulsa/ $300 per person Covers all nine classes of hazardous materials Covers shipping under IATA Certificate good for 2 years Will do group classes in local area---$3,000 plus travel costs

# Guidance for Collection and Data Entry of HIV Incidence Surveillance Information

Version 1.2

# Contents


***Introduction*** ..... ***4***

Background ..... 4

Transition to a New HIS Database ..... 4

Changes in Data Analysis ..... 5

Changes in Data Collection Forms ..... 5

Changes in the HIS Database ..... 5

Document-based Data Collection and Data Entry of TTH ..... 5

Data Collection ..... 6

Indeterminate HIV Test Results ..... 7

Data Entry ..... 7

Data Quality ..... 7

General Guidance ..... 7

Dates ..... 8

Format of this Document ..... 8

***1. Critical Data Elements for Testing & Treatment History (TTH)*** ..... ***10***

1.1 Main Source of TTH Information ..... 10

1.2 Date Patient Reported Information ..... 12

1.3 Ever had a Previous Positive HIV Test? ..... 14

1.4 Date of First Positive HIV Test ..... 17

1.5 Ever had a Negative HIV Test? ..... 19

1.6 Date of Last Negative HIV Test ..... 21

1.7 Number of Negative HIV Tests within 24 months before First Positive HIV Test ..... 24

1.8 Ever Taken Any Antiretroviral Medications (ARVs)? ..... 27

1.9 Name(s) of ARV Medication Taken ..... 30

1.10 Date ARVs First Began ..... 32

1.11 Date of Last ARV Use ..... 33

***2. Legacy Data and Other Optional TTH Data Elements*** ..... ***36***

2.1 Are You Now Taking Any ARVs? ..... 36

2.2 Ever Tested for HIV Before Today? (Legacy Pre-test form) ..... 36

2.3 Date of First HIV Test Ever ..... 37

2.4 Was the First Positive HIV Test Anonymous? ..... 37

2.5 Number of Tests 2 years before Previous Positive HIV Test (Legacy Pre-test form) ..... 38

2.6 Reason for Today's HIV Test (Legacy Pre-test form) ..... 38

2.7 Reason for First Positive HIV Test ..... 40

2.8 Name of Facility Where First Tested Positive for HIV ..... 41
2.9 State of Facility Where First Tested Positive for HIV ..... 41
2.10 Type of Facility Where First Tested Positive for HIV ..... 42
2.11 Name of Facility Where Last Tested Negative for HIV ..... 42
2.12 State of Facility Where Last Tested Negative for HIV ..... 42
2.13 Type of Facility Where First Tested Negative for HIV ..... 43

***3. IVR Database TTH Variables Not Included in eHARS ..... 43***
3.1 Date of HIV Test (reference date) ..... 43

***4. Required and Optional Data Elements for STARHS Specimen Information and Results ..... 44***
4.1 STARHS Laboratory Name ..... 44
4.2 Source Lab Specimen ID ..... 45
4.3 State Lab Specimen ID (or Other Specimen ID) ..... 46
4.4 Date of Specimen Collection (Required for HIS) ..... 46
4.5 STARHS Result Date ..... 47
4.6 Received Date ..... 47
4.7 STARHS Assay (Required for eHARS) ..... 48
4.8 Specimen Type (Recommended for HIS) ..... 49
4.9 STARHS ID (Required for eHARS and HIS) ..... 50
4.10 Standard Optical Density ..... 50
4.11 Final STARHS Result (Required for HIS) ..... 51
4.12 Reason Specimen Not Sent for STARHS ..... 52

***5. IVR Lab Variables Not Included in eHARS ..... 52***
5.1 Specimen Approved for STARHS ..... 52
5.2 State Lab ID ..... 53
5.3 HIV Diagnosis Test Type ..... 53
5.4 Results Received ..... 53

***Appendix A: Quick Reference for Data Analysts ..... 56***

***Appendix B: List of ARV Medications for TTH ..... 62***

***Appendix C: Data Entry Recommendations for HIV Incidence Surveillance Data Elements in eHARS ..... 64***

## Introduction

Since 2001, the Centers for Disease Control and Prevention (CDC) has funded a number of state and local health departments to conduct HIV incidence surveillance (HIS) through the collection of information on HIV testing and treatment history (TTH) and serologic testing algorithm for recent HIV seroconversion (STARHS) results to generate national, state and local HIV incidence estimates. Since the implementation of HIV incidence surveillance, practices have evolved as a result of advances in testing technology, availability of additional information sources and better understanding of the data necessary to estimate HIV incidence.

This document describes guidance for HIS data collection and data entry within the national HIV/AIDS reporting system. In addition to describing more specific information about the purpose and interpretation of HIS data elements, this guidance describes the wording, variable names, and available values for TTH and STARHS data in eHARS. In this document, the terms person, patient and individual are used interchangeably when referring to persons newly diagnosed with HIV.

### *Background*

The original HIS database was designed to collect information on demographics, testing and treatment history, laboratory specimens, STARHS results, and consent. Initially, STARHS used the Vironostika HIV-1 EIA (Vironostika-LS) assay for which the Food and Drug Administration (FDA) required individual consent. TTH was collected during an interview with persons seeking HIV testing (pre-test) or after receiving a new HIV diagnosis (post-test). In March 2005, the FDA allowed the use of a new assay, the BED HIV-1 Capture enzyme immunoassay (BED), for surveillance purposes only, which eliminated the need for patient consent. In 2007, CDC reduced the number data elements required for TTH and expanded the collection of TTH data elements to include provider reports, medical record review, other databases and the National HIV Monitoring and Evaluation or Program Evaluation and Monitoring System (NHM&E/PEMS). Most health departments funded for HIS used the post-test data collection form and collected data at various times, not exclusively at the time of testing.

### *Transition to a New HIS Database*

From the implementation of HIS in 2001 through most of 2011, incidence surveillance data were stored in the IVR database. In late 2011, HIS data were merged into eHARS, a document-based system used for storing all HIV surveillance information and producing datasets for local analysis and transfer to CDC. The IVR database required the entry of pre-test and post-test data, STARHS specimen information, and import of STARHS laboratory results on separate forms and tables.

In eHARS, there is one TTH document type, not separate pre-test or post-test documents. An eHARS laboratory document is used both for specimen information prior to receiving the STARHS result and STARHS results returned from the CDC STARHS laboratory. A case can have multiple TTH and laboratory documents and all data can be imported into eHARS, using All Document Import (ADI). Refer to the current *eHARS Technical Reference Guide* for more information on ADI. At conversion, pre-test and post-test information in the IVR database were converted to TTH documents in eHARS. For cases with more than one TTH document in the IVR database (e.g., a pre-test and a post-test document), separate TTH documents were stored in eHARS. Specimen information and STARHS results are now entered or imported to Laboratory documents, and HIS areas must ensure that the STARHS laboratory and any other laboratories that performed STARHS testing in the past were added by the eHARS Administrator to the local Laboratory Name list for the eHARS variable Lab Name (CLIA_ID) (See Section IV: Required and Optional Data Elements for STARHS Specimen Information and Results). At conversion, all pre-test and post-test data were converted to eHARS TTH

documents, with variable names the same as the IVR pre-test variable names, but some values may be different. Project areas have to modify previous analysis programs to reflect the new names. This guidance should be useful in making these modifications.

After conversion, all data are transferred out of eHARS to CDC, but monthly HIS datasets are required until National Data Processing (NDP) is in place. After that, quarterly datasets will be required. HIS programs will stop using the IVR database for incidence data, but some areas will continue using it for Variant, Atypical and Resistant HIV Surveillance **(**VARHS) until the incorporation of data from VARHS into eHARS.

### *Changes in Data Analysis*

Currently, TTH variable values do not appear in eHARS Person View (PV), the case-level summary of the 'best' information collected for each person that appears on the eHARS screen, but all lab documents with STARHS tests are listed. All data analyses can be performed using CDC provided programs that extract datasets from the eHARS Document datasets and Person-based dataset. Eventually, some calculated TTH variables may be included in PV and the Person-based datasets, based on a hierarchy of the best information from multiple documents. Surveillance staff will also be able to use the override option to select a different value for a particular variable from multiple documents. In the interim, CDC will provide guidance and SAS programs for selecting incidence information from multiple documents for incidence estimation and other data analysis.

Appendix A comprises a Quick Reference for Data Analysts, a guide for data managers and programmers who monitor and analyze incidence data. It provides IVR and eHARS variable names, labels, values, and formats of stored data, as well as the SAS labels that are found in eHARS document-based datasets.

### *Changes in Data Collection Forms*

This document describes changes in data collection that may require project areas to modify local data collection instruments. Surveillance areas may continue to collect data using current forms until new ones are developed. In addition, the revised adult case report form (ACRF) has a block of data elements called HIV Testing and Antiretroviral Use History that reflect the information in this guidance. Currently, information from this ACRF section must be entered in eHARS on a separate TTH document.

### *Changes in the HIS Database*

The eHARS TTH document looks different from the IVR pre-test and post-test screens. Clicking the TTH Data tab reveals first the required TTH fields (standard data elements described in *Technical Guidance for HIV/AIDS Surveillance Programs*), followed by the optional/legacy fields. The data fields have simple, short labels. A few fields have changes in the values that can be selected from drop-down menus. Some values, such as '99/99/9999' for a missing date, are not accepted in eHARS.

Specimen information and STARHS results are captured on an eHARS laboratory document, separate from the laboratory document for the diagnostic HIV test. The names of specimen variables in eHARS differ from the names in the IVR database, but the main data elements are captured in eHARS. Some data elements such as *Specimen Approved for STARHS* (LAPPRVE) and *Were Results Received* (RESULTSRECEIVED) do not appear in eHARS.

### *Document-based Data Collection and Data Entry of TTH*

Document-based data collection and entry differ somewhat from past practice. Using the IVR database, surveillance areas generally selected one form type (pre-test or post-test) for entering TTH information. If additional data were collected in another investigation from a second source, some

surveillance areas recorded those data on the other form, but many just updated the original document. As a result, HIS data forms contained mixed and sometimes conflicting data.

Surveillance data are often collected by non-incidence surveillance staff such as disease intervention specialists (DIS) during client interviews, field staff conducting medical chart reviews, or medical providers filling out case report forms. These collectors usually fill out one form but may collect information from more than one source. For example, an interviewer at a sexually-transmitted disease (STD) clinic may use information from a database for *Date of Last Negative HIV Test* to supplement patient-reported information.

To be consistent with the principles of document-based data entry, new eHARS documents should be used to enter data from multiple forms, from multiple sources, or to update information. At the top of the TTH document screen, a new field, *Main Source of TTH Information* has replaced the IVR database variable, *Was the Testing History Questionnaire Implemented?*, which indicated whether a patient interview had been conducted. *Main Source of TTH Information* is important for document-based HIS because it allows CDC and the project area to track the relative proportions of TTH documents collected through various reporting methods (e.g., patient interview, medical record review, passive case reporting) for the purpose of program improvement.

Another change is that eHARS also has additional tabs associated with every document. For guidance on recommended fields that need to be completed on other eHARS tabs (Form Info, Identification, and Demographics tabs), consult Appendix C, *Data Entry Requirements for HIV Incidence Surveillance Data Elements in eHARS.*

### Data Collection

The general principle in data collection is to collect the most accurate information available to characterize a person's HIV testing and treatment history. When investigating a case or completing a case report form, data collection staff and medical providers may have access to multiple sources of information, including the patient, medical records, laboratory reports, and other databases. It is important to examine all the information on hand and ensure that the data recorded on the TTH form most accurately reflect the patient's actual HIV testing history and antiretroviral (ARV) use. The purpose of the TTH is not to study the patient's knowledge of his or her HIV status, but rather, (a) to determine testing frequency in order to estimate the probability of being tested during the BED recency period, and (b) to determine whether ARV use might have affected BED results.

Incidence surveillance staff should record data collected during separate investigations on separate TTH data collection forms. A case report form sent in by a provider and a data collection form filled out by field staff during a medical chart review each have a different *Main Source of TTH Information* and should be recorded on different forms, even if both collectors used an ACRF.

Sometimes there will be mixed sources during one investigation and the data collector is filling out one form. A data abstractor conducting a medical chart review should represent the best information obtained from all sources reviewed, whether from patient self-report found in physician notes or laboratory results noted in the chart. Similarly, for a provider report, the information will reflect all the information available, including the provider's knowledge of a previous positive or negative test that was not reported by the patient. However, when an interviewer collects most of the TTH during the patient interview and obtains the rest of the information from a medical chart or database, two forms should be completed because these represent two different data collection methods. For example, during an interview, a patient reports never having had a negative HIV test, but the interviewer also finds in a database the date of a previous negative test for which the patient never returned for the results. The interviewer should enter the information from the database on a separate TTH form. If

there are two separate investigations (e.g., an interview and a complete chart review, or two medical record reviews at different locations or time), two TTH forms should be completed. A TTH form completed during investigation should summarize all the information obtained during that particular investigation. If an additional TTH form is received by the health department with data identical to a previously entered TTH, it is not necessary to create an additional TTH document. However, surveillance staff must check to see that all elements are identical to the original document.

#### INDETERMINATE HIV TEST RESULTS

Data collectors should ignore indeterminate test results in recording responses to previous testing history questions, including previous positive tests, previous negative tests, and number of negative tests. An indeterminate test is neither positive nor negative for HIV. The guidance for indeterminate tests differs from the guidance for unknown test results, (e.g., where patient never received test results) because it may be possible to determine from another source that the unknown test result was either positive or negative.

### *Data Entry*

Data entry staff should enter data in eHARS as reported on the data collection form. Incidence data may be entered manually or imported into eHARS. Incidence surveillance staff should always enter data from different data collection instruments received on separate TTH documents; for example, a provider case report form and a form filled out by field staff from a medical chart review. Each has a different *Main Source of TTH Information*. If medical record reviews were conducted at two different locations, the data should have been recorded on 2 separate TTH document forms and should subsequently be entered on separate TTH documents in eHARS. Each eHARS document also may have a different *eHARS Document Source* – reflecting the type of setting where the data were collected – entered on the eHARS Form Info tab. For more information on recommended fields to be completed on eHARS documents, refer to **Appendix C.**

When a data conflict is found on a single data collection form, data entry staff should not make changes to the entries. Data entry staff should follow program procedures to resolve the data conflicts which may include consultation with the Incidence Surveillance Coordinator or follow-up with the individual who collected the data.

### *Data Quality*

Sometimes, due to the nature of the data collection process (e.g., different sources) data conflicts may arise. Each surveillance area should have Standard Operating Procedures (SOP) for resolving data conflicts on data collection forms. The Incidence Surveillance Coordinator should investigate the conflict and verify the information collected to determine if any further action is needed. Further actions may include entering another document with the correct information or overriding a field in eHARS Person View when incidence data are included in the person view document. Moreover, CDC recommends that surveillance areas conduct regular data checks and data cleaning activities.

### *General Guidance*

The following are descriptions of the common response options:

- **"Yes"** indicates that there was sufficient evidence that the event occurred. Evidence can be from patient self-report, health care provider note, or laboratory documentation.
- **"No"** indicates that there was sufficient evidence that the event did not occur. Evidence can be from patient self-report or health care provider documentation that definitively states that the

given event never occurred (e.g., no previous positive test, never had a negative test, no ARV use).

- **"Don't know"** indicates that the patient reported "don't know," the health care provider documented "unknown," or there was insufficient evidence for or against (supporting or denying) the occurrence of the event.
- **Blank** indicates that the usual data sources were not investigated and/or the health care provider/staff were not asked.
- **"Refused"** means patient refused, health care provider recorded "refused," or facility refused to permit the medical record review.

### *Dates*

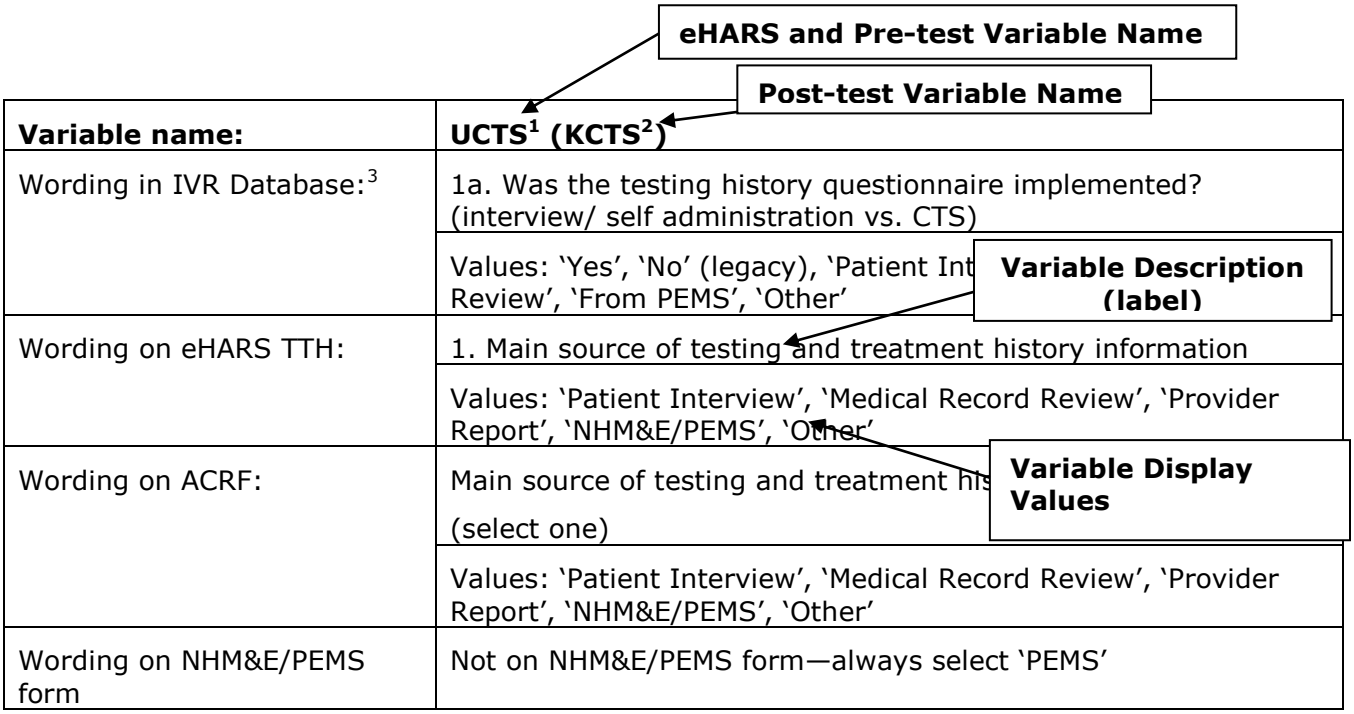
Dates are very important data elements for HIS. The date of first positive HIV test, date of last negative HIV test, and dates of ARV use help to classify persons as new or repeat testers and as having recent or long-term infections. In the hierarchy of incidence estimation algorithms, dates supersede 'yes' and 'no' answers.

For all TTH dates, eHARS has space for month, day, and year. For most dates, the IVR database and data collection forms only contained month and year. In many cases, the day or even the month will be unknown. It is acceptable to leave day or month blank (entering '..' for missing month or day in eHARS, such as ../../2001), but it is important to try to obtain at least the year, which can be used in calculations. An approximate date reported by a patient is better than no date.

### *Format of this Document*

This document provides the following for each variable used for HIS:

- Name(s) of variable (eHARS TTH, IVR Pre-test and Post-test)
- IVR, eHARS, and NHM&E/PEMS labels
- Available values
- Description
- Purpose of the variable
- History (if applicable)
- Possible sources for the information
- Specific information on data collection
- Specific instructions for data entry
- Examples by source of information




| Variable name: | UCTS[1] (KCTS[2]) |
|---|---|
| Wording in IVR Database:[3] | 1a. Was the testing history questionnaire implemented? (interview/ self administration vs. CTS) |
| | Values: 'Yes', 'No' (legacy), 'Patient Int Review', 'From PEMS', 'Other' |
| Wording on eHARS TTH: | 1. Main source of testing and treatment history information |
| | Values: 'Patient Interview', 'Medical Record Review', 'Provider Report', 'NHM&E/PEMS', 'Other' |
| Wording on ACRF: | Main source of testing and treatment his<br>(select one) |
| | Values: 'Patient Interview', 'Medical Record Review', 'Provider Report', 'NHM&E/PEMS', 'Other' |
| Wording on NHM&E/PEMS form | Not on NHM&E/PEMS form—always select 'PEMS' |

[1] eHARS variable name, which is the Pre-test variable name

[2] Post-test variable name (no longer used in eHARS)

[3] From Pre-test version in IVR database

Examples of data collection/entry concerns are presented by source of TTH information because there are often different instructions with each type of source. This document provides guidance to HIV Incidence Surveillance Coordinators who will train staff or providers in the collection and data entry of TTH using the revised adult case report form (ACRF) and other local data collection forms.

Certain formatting conventions are used. Refer to the example below.

- Descriptions of variables appear in italics (e.g., *Main Source of TTH Information*).
- Specific values for variables appear in single quote marks (e.g., 'patient interview') and refer to the display values, not necessarily the values stored in databases.
- Variable names appear in ALL CAPS (e.g., UCTS).
- eHARS variable names for fields on the TTH document appear in ALL CAPS first and correspond to former Pre-test variables, with retired (mostly former Post-test) variables in parentheses, such as UCTS (KCTS).
- For STARHS specimen information and test results, eHARS variable names appear in ALL CAPS with IVR Access database variable names in parentheses, such as LOINC_CD (ASSAY).

# 1. Critical Data Elements for Testing & Treatment History (TTH)

## *1.1 TTH Data Element #1: Main Source of TTH Information*

| **Variable name:** | **UCTS[1] (KCTS[2])** |
|---|---|
| Wording in IVR Database:[3] | 1a. Was the testing history questionnaire implemented? (interview/ self administration vs. CTS) |
| | Values: 'Yes', 'No' (legacy), 'Patient Interview', 'Medical Record Review', 'From PEMS', 'Other' |
| Wording on eHARS TTH: | 1. Main source of testing and treatment history information |
| | Values: '1-Provider Report', '2-Patient Interview', '3-Medical Record Review', '4- NHM&E/PEMS', '5-Other' |
| Wording on ACRF: | Main source of testing and treatment history information (select one) |
| | Values: 'Patient Interview', 'Medical Record Review', 'Provider Report', 'NHM&E/PEMS', 'Other' |
| Wording on NHM&E/PEMS form | Not on NHM&E/PEMS form—always select 'NHM&E/PEMS' |

[1] eHARS variable name, which is the Pre-test variable name

[2] Former Post-test variable name in parenthesis

[3] From Pre-test version in IVR database

### *Description*

The *Main Source of TTH Information* variable reports the method for obtaining most of the information on the TTH document. This variable is different from the eHARS *Document Source* variable which describes the type of provider or facility where the case information was gathered (e.g., hospital, laboratory, physician's office, HIV clinic, or other database).

### *Purpose*

The purpose of this variable is to characterize the source of the majority of data elements for a given TTH document. CDC and surveillance areas will use this variable to evaluate the amount of information that comes from various data collection sources which will inform incidence estimation and program improvement. Also, it will be useful for data quality purposes to know if the source was patient interview or another source when investigating conflicting information within a document or across documents.

### *History*

In the original data collection forms, this question was worded "*Was the testing history questionnaire implemented?*" (yes/no). This caused confusion, so more categories ('Patient Interview', 'Medical Record Review', 'From PEMS', 'Other'*)* were added in the revised IVR Access database deployed in July 2006. This variable was not included in the Standard Data Elements after 2007; however, it became important for incidence estimation, program improvement, and data quality starting in 2010. Surveillance areas currently should be using this variable. In the past, the IVR database had no specific category to designate Provider Report, so data entry staff selected 'other'. After eHARS conversion, the category 'provider report' became available.

### *Descriptions of Sources*

Patient Interview: Patient was directly asked most of these questions by a disease intervention specialist (DIS), health department (HD) staff, or a health care provider who asked the series of TTH questions from a structured TTH form and who has received some training in the proper collection of TTH information. If a chart review occurs in addition to the interview, two forms should be filled out, one with 'interview' and one with 'medical record review' as the *Main Source of TTH Information.*

Medical Record Review: HD staff obtains the information through review of medical charts or electronic medical records or databases. A few elements may come from the provider or other source obtained as part of a medical record abstraction investigation. When HD staff visits a health care provider's office and extracts data from medical records, the source is 'medical record review', even if the data collection form is the surveillance area's provider case report form.

Provider Report: Information was obtained from a case report form submitted by a health care provider or from a phone call with a provider. The health care provider does not indicate having directly asked patients for most of this information from a TTH form. Information was found by the provider from chart notes, laboratory reports, or recollection from discussion with the patient or another health care provider. If the provider administered a series of questions from an ACRF or other form, then 'patient interview' could be selected, depending on local policies (see note under **Data Collection** below).

NHM&E/PEMS: The Program Evaluation and Monitoring System (PEMS) is also known as National HIV Monitoring and Evaluation (NHM&E) and the proposed test form is called the NHM&E HIV Test Form. Information was obtained from NHM&E or earlier PEMS forms (e.g., Parts 1 and 3), locally developed CTR forms, or imported from a NHM&E/PEMS database. It is understood that most of NHM&E/PEMS data come from patient interview but this source selection clarifies the source as NHM&E/PEMS, which is useful for evaluation. For this guidance, we use the two names, PEMS and NHM&E, interchangeably.

Other: Information was obtained from another source. If the only source is a database, the ADAP database for example, select 'other'.

If the source of data is unknown, then leave *Main Source of TTH Information* blank.

### *Data Collection*

The data collectors should select the category that best describes the source of the TTH data elements on this form. If there are data elements collected from two or more of the sources described above (e.g., an interview and a chart review), then the data collector should use two separate forms, in accordance with document-based data collection principles.

**Note**: Each local surveillance area should consider all of the regular sources of information that contribute to the TTH and assign a value for *Main Source of TTH Information* for each type of source. Even if a provider checks 'patient interview' on the ACRF, the value should be entered as 'provider report' unless there is evidence that the provider conducted a structured interview using specific TTH questions. This is because the questions may not be asked or interpreted exactly as this Guidance document describes. Therefore, Incidence Surveillance Coordinators in consultation with the CDC epidemiologist may need to make project area-specific decisions on how to classify information from certain local forms or particular reporting sources. For example, if there are providers at an HIV clinic that have received training on the collection of TTH, then the Incidence Surveillance Coordinator may decide that forms received from that clinic should be designated 'patient interview'.

### *Data Entry*

If the data source does not appear on the data collection form, the data entry person, in consultation with the Incidence Surveillance Coordinator, should enter the best answer that characterizes the data source. For example, if the data come from an ACRF or a local provider case report form sent to the health department by a provider, enter 'provider report'. If the information comes from a data collection form that field staff uses to conduct chart abstraction, even if it is the ACRF, enter 'medical record review'. If the form is a known data collection instrument from a STD clinic interview, select 'patient interview'. If the data are on a NHM&E/PEMS form or from a NHM&E/PEMS database, enter 'NHM&E/PEMS'. The data entry person should consult with the Incidence Surveillance Coordinator if uncertain about the data source.

## *1.2 TTH Data Element #2: Date Patient Reported Information*

| **Variable name and format[1]:** | **UQINTD (KQINTD) mm/dd/yyyy** |
|---|---|
| Wording in IVR Database: | 1. Today's Date (Use THQ Interviewing Date) |
| Wording on eHARS TTH: | 2. Date patient reported information |
| Wording on ACRF: | Date patient reported information |
| Wording on NHM&E/PEMS form: | Session Date (Part 1) or Date Information Collected (Part 3) |
| Wording on proposed NHM&E form: | Date Client Reported Information (Part 3) |

[1] Format displayed on eHARS data entry screen

### *Description*

The *Date Patient Reported Information* variable represents different dates, depending on the circumstances in which the information was primarily obtained. This variable represents the date of the patient interview if information was obtained face-to-face, or the date of the provider note for the last patient encounter at which TTH information was ascertained from a medical record. If there was no patient contact, then the date of receipt of TTH information from a laboratory report or database should be used. If none of these dates are available, date of medical record review or date provider completed the report form should be used.

### *Purpose*

This date may be used in the future for some calculations in analyses and to select the best incidence values for eHARS Person View (PV) when there are multiple TTH documents. In the eHARS PV hierarchy, some TTH values may be selected from the earliest TTH document based on this date or, if the date is blank, the date the document was entered. However, individual variables can be overridden in PV by the Incidence Surveillance Coordinator to best represent a case's testing and treatment history.

### *Sources*

Sources for this variable could include interview date on a NHM&E, CTS or other patient interview form, last date of a patient visit noted in a medical chart from which the TTH information was ascertained, date of interview in the NHM&E/PEMS database, date reported on a provider report form,

and, if *Date Patient Reported Information* was blank, date health care provider filled out a case report form or date received at the health department.

### *Data Collection*

In general, the data collector should select the last date of interview or patient visit that contributed to the testing and treatment information. It represents the point in time when the TTH information was obtained from a patient. It does not necessarily mean the last patient visit. If data only were obtained from a laboratory report or from a database, the date should be the date when the data were received by the provider or health department. If none of these dates are available, use the date when data were collected (i.e., the date of medical record review or date provider completed the case report form). If those dates are missing, use the date the information was received by the health department.

### *Data Entry*

If there are two different data collection forms or sources, the data entry person should enter them as separate TTH documents, each with the appropriate *Date Patient Reported Information.* If the date is missing, enter the date the form was received at the health department.

### *Examples by Source of Information:*

Patient Interview: Date of interview should be entered. An interview can be conducted by a health care provider, Disease Intervention Specialist (DIS), or Health Department (HD) staff using the standard TTH questions.

- More than one interview was conducted—interviewer(s) should complete separate TTH documents, one for each interview date.
- Interview was conducted on two different days, due to an interruption. Interviewer should enter all of the information on one form and select the last date when interview was completed.

Medical Record Review: Use the last date of the patient encounter or provider's note that contributed to the TTH information for this form. Do not use the actual medical record review date unless there is no other date to use.

- Medical record review was conducted at two different locations—abstractor should enter two separate TTH documents with appropriate dates.
- Medical record review occurred at two different points in time—abstractor(s) should enter separate TTH documents with the appropriate dates.
- Chart abstractor started a chart review one day and completed the review on the same form a week later—abstractor should record the date of the last patient encounter with TTH information.
- Chart abstractor also found a pharmacy record for ARVs dated after the last patient encounter—abstractor should record the date of the last patient encounter with TTH information for *Date Patient Reported Information* because it reflects the time when the patient answered questions. He should use the pharmacy information for recording dates of ARV use.
- Chart abstractor found no TTH information in the provider notes for patient visits, but found a negative HIV test date in a clinic testing database—she should enter the date the test result was received. If there is no received date, then enter date of chart abstraction.

Provider Report: Providers should be instructed to record on the provider report form the date when most of this information was obtained from the patient or another data source, or, lacking that, the date when the provider completed the report.

- Provider filled out a case report form—provider should record the date of the last patient encounter that provided TTH information.
- Provider filled out a case report form for a person with a positive HIV test result but there was no patient encounter—provider should enter date that the form was filled out.
- Information was only gathered from laboratory reports—provider should enter the date on which the provider received the laboratory results or, if unavailable, the date he or she filled out the form.
- The last encounter with patient where TTH questions were answered was in April 2010, but the pharmacy record shows a prescription for ARVs on October 2010—record the date of the patient encounter, April 2010.
- The only information came from pharmacy records—report the date the prescription was filled or, if missing, the date the provider form was completed.
- HIS staff member called the provider because there was no TTH information on the ACRF submitted to the health department. If the provider does not have any TTH information there are 2 choices: 1) do not complete a TTH form as no new information was provided, or 2) complete a TTH form by entering 'don't know' for several variables based on the phone call. In this case, enter the date of the phone call for *Date Patient Reported Information*.
- No reported *Date Patient Reported Information* on ACRF or provider report form—data entry person should enter the date the provider completed the form, if available, or when the case report was received by the HD.

NHM&E/PEMS*:* HIS data entry person should enter the date of interview obtained from NHM&E, PEMS or CTS form or database.

Do not leave this date blank. If *Date Patient Reported Information* is unknown, enter the date when the document was received at the health department. This date may be used for prioritizing the selection of incidence data in eHARS PV when there are multiple TTH documents.

### *1.3 TTH Data Element #3: Ever had a Previous Positive HIV Test?*

| **Variable name and format:** | **UPASTP[1] Yes/No/Refused/Don't know** |
|---|---|
| Wording in IVR Database (only on Pre-test form): | 4. Have you ever had a positive HIV test result? |
| Wording on eHARS TTH: | 3. Ever had a previous positive HIV test? |
| | Values: 'N – No', 'Y – Yes', 'R – Refused', 'D – Do not know' |
| Wording on ACRF: | Ever had previous positive HIV test? |
| Wording on PEMS form: | Part 1, Previous HIV Test? (#26), Self-reported Result (#27), and Date of Last Test (#28/29) |
| Wording on proposed NHM&E form: | Has client ever had a previous positive HIV test? (Part 3) |
| | Values: 'Yes', 'No', 'Don't know', 'Declined' |

[1] eHARS and Pre-test variable name

### *Description*

The *Ever Had a Previous Positive HIV Test* variable provides information about the patient ever having a positive HIV test before the current test. This variable originally only appeared on the Pre-test TTH form but was added to the CDC ACRF in 2010 to clarify information on the first positive HIV test. Surveillance areas should make sure this variable is added to local data collection forms.

### *Purpose*

The purpose of this variable is to ascertain whether a positive HIV test occurred earlier than the eHARS calculated *HIV Disease Diagnosis Date* (HIV_AIDS_DX_DT), but was not reported to the HIV surveillance system; for example, a patient could have been diagnosed in another state/country or tested anonymously. Having an earlier test, accompanied by an approximate date, might indicate a long-term HIV infection. This information will also be used for TTH completeness reports starting in 2011 and to evaluate the validity of the variable, *Date of First Positive HIV Test*.

In the absence of an earlier, self-reported *Date of First Positive HIV Test* (#4), CDC and local analytical programs assume that the HIV disease diagnosis date is the collection date of specimen used for the first positive HIV test. The variable allows the entry of a 'no' to confirm that the person reported never having a previous positive test. We can use this to learn for what proportion of cases it is known that this is the first positive test.

### *Sources*

This information may come from patient self-report, evidence of an anonymous test, record of test performed in another country, a doctor's note, discussion with another provider, or other source that is probably not a laboratory report. A confirmed positive from an HIV home test kit should be considered the same as a self-reported anonymous test. It is not reported to the health department.

### *Data Collection*

Data collection may differ by source of information. In a patient interview, a person will likely state that he or she has or has not had a previous positive test. Persons with a previous positive test often were tested anonymously and are now converting to confidential testing. Also, the person may have been tested in another state or was tested previously and is now seeking care. If so, record 'yes' and make an effort to obtain the date of first positive for *Date of First Positive HIV Test* (variable #4). Ignore indeterminate and false positive tests. If the only previous HIV test was indeterminate, do not consider that test as a positive test; in that case, enter 'no'. For medical record review or provider report, it is more difficult to ascertain whether there was a previous positive HIV test unless the patient was specifically asked and the information was recorded in the medical notes. If there is no evidence either way, enter 'Don't know' on the data collection form. This is useful information for interpreting a date for the *Date of First Positive HIV Test* (variable #4 below).

### *Data Entry*

The data entry person should enter the response recorded on the data collection form for this question. Otherwise, leave the field blank. If the surveillance area does not have this field on current forms, it should be added to TTH collection forms. If a conflict is identified, the data entry person should follow standard operating procedures (SOP) for resolving data conflicts and consult with the Incidence Surveillance Coordinator.

### *Examples by Source of Information:*

Patient Interview:

- Patient self-reported a previous test—interviewer should record 'yes'.
- Patient reported a positive HIV test from a home test kit that was sent into a laboratory—interviewer should record 'yes'.
- Only previous test was 'indeterminate'—interviewer should record 'no'.
- Patient reported a positive oral screening test followed by a negative Western Blot—interviewer should record 'no'. This test is considered a negative HIV test.
- Patient had a false positive HIV test when she was pregnant 3 years ago and was confirmed to be HIV negative 6 months later—record 'no'.
- Patient did not know—interviewer should record 'don't know'.
- Patient never got his results—interviewer should record 'don't know'.
- Patient reported 'no' or 'don't know' but the interviewer later found evidence of an earlier positive test– interviewer should leave the patient's answer as 'no'. The date of the earlier positive test should be entered on a separate TTH document. If it is a documented positive confirmatory test, it also should be entered on an ACRF or Laboratory document.

Medical Record Review:

- Chart abstractor found evidence/report of a previous positive test in the medical record; for example, a doctor's note that said 'patient tested in CA in 2006'—abstractor should record 'yes'.
- Chart abstractor found information about a previous test in a clinic database, regardless of whether patient knew about it—abstractor should record 'yes'.
- Medical record indicated that the patient never had a previous positive test; for example, a note saying 'this is the patient's first positive HIV test'—abstractor should record 'no'.
- The patient reported that the only previous HIV test patient had was 'indeterminate' —abstractor should record 'no'.
- Patient reported she never got her results—abstractor should record 'don't know'.
- There was no definitive evidence either way; for example, the provider did not ask the question or made no note in the chart—abstractor should record 'don't know'. This indicates that effort was made to find the answer but it was unavailable. Do not assume that no evidence is the same as no previous positive test.

Provider Report: Health care providers should be encouraged to ask patients about previous testing and record notes in the medical record.

- Provider knows from patient report that this was the first positive HIV test that the patient ever received—provider should record 'no'.
- Provider knows there was an earlier positive HIV test before the one being reported, but has no date—provider should record 'yes'.
- Provider does not know about any previous positive HIV test—provider should record 'don't know'.

NHM&E/PEMS: The HIV Surveillance data entry person or import program should calculate from items 26, 27, 28, and 29 from Part 1 to see if there was a previous positive HIV test, or record the

answer from the TTH question on Part 3 of the new NHM&E form. Enter the information obtained from a NHM&E, PEMS or CTS form or database. If there is no information, leave blank.

If there was no patient interview, no medical chart review, no provider report, and no other follow-up, the answer should be left blank.

### *1.4 TTH Data Element #4: Date of First Positive HIV Test*

| **Variable name and format:** | **UFPOSD (KFPOSD) mm/dd/yyyy** |
|---|---|
| Wording in IVR Database: | 4b. What was the month/year of the very first time you ever tested positive for HIV? List when you got the test, not when you got the results. |
| Wording on eHARS TTH: | 4. Date of first positive HIV test |
| Wording on ACRF: | Date of first positive HIV test |
| Wording on NHM&E/PEMS form: | Date of first positive HIV test (Part 3) |

#### *Description*

The *Date of First Positive HIV Test* variable is the date of the earliest known positive HIV test for the patient. It represents the date that the specimen was collected for the very first positive HIV test. This date could represent an anonymous test that will never be reported to the national HIV surveillance system. Most of the time, this date is self-reported.

**Note**: Any documented, positive HIV laboratory test result that is earlier than the current calculated eHARS *HIV Disease Diagnosis Date* (HIV_AIDS_DX_DT) should be entered in eHARS on an ACRF or on a Laboratory document. That will result in the recalculation of HIV_AIDS_DX_DT in eHARS Person View.

#### *Purpose*

This date is important for identifying cases that tested positive for HIV earlier than the eHARS calculated *HIV Disease Diagnosis Date* (HIV_AIDS_DX_DT) and therefore may not represent new diagnoses. For most data analyses, the eHARS *HIV Disease Diagnosis Date* is used. However, during incidence estimation, if the TTH self-reported *Date of First Positive HIV Test* is earlier than the eHARS date, the TTH date is used for reclassifying BED results and for calculating the inter-test interval (T). If the self-reported *Date of First Positive HIV Test* is more than 6 months earlier than the eHARS diagnosis date, BED results are set to 'long-term'; if the *Date of First Positive HIV Test* is 1–6 months before the eHARS date, BED is set to missing and results are imputed. If the *Date of First Positive HIV Test* is earlier than eHARS diagnosis date, it will be used in the above calculation, even if *Ever had a Previous Positive HIV Test* is 'no'. If *Date of First Positive HIV Test* is blank, it is assumed, for analysis purposes, to be the same as the *HIV Disease Diagnosis Date.*

#### *Sources*

Sources for this variable could be patient self-report, report of an anonymous test, report of a test performed in another jurisdiction, record of a test in a clinic database, doctor's note, discussion with another provider, or other source.

### *Data Collection*

The self-reported *Date of First Positive HIV Test* could be the same as the documented date of the confirmatory HIV test report or eHARS calculated HIV disease diagnosis date. During interviews, chart abstraction, or completion of the ACRF, some data collectors will enter the date of the current HIV positive test for the purposes of calculating subsequent answers. When it is known that the current test date is the date of first positive HIV test, it is important to answer 'no' for the *Ever had a Previous Positive HIV Test* data element (#3 above). That will clarify why the date is the same as the eHARS HIV diagnosis date. Do not record any dates for 'indeterminate' HIV tests, false positive tests or for tests with unknown results.

### *Data Entry*

Data entry person should enter the date recorded on the form. If a conflict is identified, follow standard operating procedures for resolving data conflicts and consult with the Incidence Surveillance Coordinator.

### *Examples by Source of Information:*

Patient Interview*:*

- Patient reported having a previous positive HIV test (e.g., an anonymous test or a test in another state)—interviewer should record the date of the first one. The interviewer can use prompts to assist the patient to remember the date, at least the year of the test.
- Patient stated she has never had a previous positive test—record current test date and select 'no' for question #3.
- Patient said he never had a previous positive, but the interviewer has information that there was a previous positive test for which the patient never returned to get the results—interviewer should enter the date of the previous positive test. Interviewer may share the test results in accordance with local protocols and HIPAA policies.
- After the interview, the interviewer found evidence of a previous positive test—enter the date on a separate TTH document.
- Patient did not know if she ever had a positive HIV test. Enter 'don't know' for *Ever had a Previous Positive HIV Test* (#3) and leave *Date of First Positive HIV Test* blank.

Medical Record Review:

- Medical record abstractor is not sure what the eHARS date of HIV infection is—abstractor should enter the date of the earliest known HIV positive test noted in the record. This would be the date the specimen was collected for the initial HIV test.
- Medical record, including self-report, indicated the patient was diagnosed or tested at earlier time than the current test being investigated—abstractor should record the date noted in the record and enter 'yes' for *Ever had a Previous Positive HIV Test* (#3).
- Medical record indicates the patient was never tested before—abstractor should record 'no' for *Ever had a Previous Positive HIV Test* (#3) and record the date of the first known positive test. This should be the date of the current test.
- There is no information in the chart about a previous test—abstractor should record 'don't know' for *Ever had a Previous Positive HIV Test* (#3) and leave *Date of First Positive HIV Test* blank.

Provider Report: Often health care providers enter the current test date when completing a provider case report form.

- Patient reported a previous test date to the provider or the provider has access to information about a previous positive—provider should record the date and enter 'yes' for *Ever had a Previous Positive HIV Test* (#3).
- Patient reported to the provider that this was his first HIV test ever—provider enters 'No' for *Ever had a Previous Positive HIV Test* (#3) and records the date of current test.
- Provider doesn't know if there was a previous positive—provider should record 'don't know' for *Ever had a Previous Positive HIV Test* (#3) and leave *Date of First Positive HIV Test* blank.

NHM&E/PEMS: The HIV Surveillance data entry person should enter the information obtained from NHM&E/PEMS (Part 3) or CTS form or database. If there is no information, leave blank.

If the value entered on any form or into any database is '999999' or there was no patient interview, no medical chart review, no provider report, and no other follow-up, leave the *Date of First Positive HIV Test* blank.

### *1.5 TTH Data Element #5: Ever had a Negative HIV Test?*

| **Variable name and format:** | **UNGTST (KNGTST) Yes/No/Refused/Don't know** |
|---|---|
| Wording in IVR Database: | 4f. Have you ever had an HIV test that was negative? |
| Wording on eHARS TTH | 5. Ever had a negative HIV test? |
| | Values: 'N – No', 'Y – Yes', 'R – Refused', 'D – Do not know' |
| Wording on ACRF | Ever tested HIV negative? |
| Wording on NHM&E/PEMS form: | Has client ever had a negative HIV test? (Part 3) |
| | Values: 'Yes', 'No', 'Don't know', 'Declined' |

#### *Description*

This variable, *Ever Had a Negative HIV Test,* captures whether or not the person ever had a negative HIV test result at any time in the past. In the initial IVR database, this variable appeared only on the Post-test form but was added to the Pre-test form with the revised IVR Access database, version 3.1 released in July 2006.

#### *Purpose*

This variable is one of three data elements (variables #5, 6, and 7) used to classify cases as new testers or repeat testers. This distinction is important for incidence estimation because the probability of being classified as recent by the BED assay is calculated separately for new testers and repeat testers. Persons with a 'yes' answer—indicating a previous negative HIV test—are classified as repeat testers, and those with a 'no' answer are classified as new testers. For the purposes of incidence estimation, when *Ever Had a Negative HIV Test* is unknown or missing, *Date of Last Negative Test* is blank, and *Number of Negative HIV Tests Within 24 Months before First Positive HIV Test* is unknown, missing, or '0' (zero), persons will be classified as a new or repeat tester using multiple imputation. Multiple imputation is a statistical process that will assign cases to either the 'new' or 'repeat' tester groups.

### *Sources*

Sources for this variable could be patient self-report, laboratory report of a negative HIV test, record of a test in a clinic database, doctor's note, discussion with another provider, or other source.

### *Data Collection*

Since this data element is used to classify cases as new testers or repeat testers for incidence estimation, it is important to have accurate information because misclassification can strongly impact the accuracy of the incidence estimate. The data collector should not make the assumption that there was never a negative test when he or she could not find any information about a negative test. Because people are tested for HIV in many different venues or with different providers, the absence of information about previous testing does not mean that previous tests did not occur.

If an interviewer has knowledge of a previous negative HIV test, he or she should prompt the patient to recall it. If the interviewer finds evidence of a previous negative test after the interview, even though the patient answered 'no', the interviewer should record the *Date of the Last Negative HIV Test* (#6) but he or she does not need to change the patient's response. The algorithm for classifying new testers and repeat testers gives priority to the *Date of Last Negative HIV Test* and will assign the person to the repeat tester category. It is best practice for the data collector to enter the date of the negative test on a separate TTH form when those data were found separately from other information.

If the patient reports that he or she doesn't know, if the health care provider does not know whether the patient had any negative HIV tests, or there was insufficient documented evidence in the medical record supporting or denying the occurrence of a negative test, the data collector should select 'don't know'. 'Don't know' and 'can't find evidence' are treated similarly. The field should be left blank if the usual data sources have not been investigated and/or the health care provider did not ask the patient.

If the patient's only previous HIV test was positive, select 'no'. Ignore indeterminate tests. If the only previous test was indeterminate, select 'no', because there is no evidence of a previous negative. Similarly, an undetectable viral load is not evidence of a negative HIV test.

### *Data Entry*

The data entry person should enter the data as recorded on the form. If the data entry person identifies a conflict, he or she should consult with the Incidence Surveillance Coordinator and follow the process outlined in the SOP. If there is no information on the form, leave blank.

### *Examples by Source of Information:*

<u>Patient Interview</u>:

- Patient was asked this question—interviewer should record the answer as given.
- Patient reported having a negative test result from a home HIV test kit that was sent to a laboratory for testing—interviewer should record 'yes'.
- Patient was tested but did not know the test results—interviewer should record 'don't know'
- Patient reported 'no' or 'don't know' to *Ever Had a Negative HIV Test* but interviewer found evidence of a previous negative HIV test before or during interview—interviewer may prompt the patient by sharing the previous negative test results in accordance with local protocols and HIPAA policies. Regardless of patient response, record the *Date of the Last Negative Test* (#6, below).

- Patient answered "no," but interviewer later finds evidence of a previous negative HIV test, interviewer should not change the response but record the *Date of the Last Negative Test* (#6, below). The date should be recorded on a separate form, if possible.
- Patient was unsure about whether he had a negative test—interviewer should record 'don't know'.

Medical Record Review: Chart abstractors should examine the medical record for a provider note or laboratory evidence of a negative test.

- There was a report of a negative HIV test, with or without a date—abstractor should enter 'yes'.
- There was a note indicating that the patient had never tested before—abstractor should record 'no'.
- There was no note or other indication that the patient ever had a negative test or not—abstractor should record 'don't know'. This indicates that effort was made to find the answer but it was unavailable.

Provider Report: Providers should be encouraged to ask patients about previous HIV tests.

- Patient reported having a previous negative HIV test (with or without a date)—provider should report 'yes'.
- Provider has access to information about at least one previous negative HIV test—provider should report 'yes'.
- Patient indicates she never tested before or she never had a negative HIV test—provider should report 'no'.
- Provider finds no evidence of a previous negative test in her notes or patient reports 'don't know'—provider should report 'don't know'.

NHM&E/PEMS: Data entry person should enter the information obtained from NHM&E/PEMS (part 3) or CTS form or database. If there is no information, leave blank.

If there was no patient interview, no medical chart review, no provider report, or no other follow-up, the answer should be left blank.

### *1.6 TTH Data Element #6: Date of Last Negative HIV Test*

| **Variable name and format:** | **ULSTND (KLSTND) mm/dd/yyyy** |
|---|---|
| Wording in IVR Database: | 4f. What was the month and year that you got your last negative HIV test? List when you got the test, not when you got the results. |
| Wording on eHARS TTH | 6. Date of last negative HIV test |
| Wording on ACRF-TTH | Date of last negative HIV test |
| Wording on ACRF-LAB | Date of last documented negative HIV test |
| Wording on NHM&E/PEMS Form | Date of last negative HIV test (Part 3) |

### *Description*

This variable, *Date of Last Negative HIV Test*, is the date of the last known negative HIV test and represents a point in time when the person was known not to be infected with HIV.

### *Purpose*

This is one of the most important data elements for incidence estimation. It is used to classify cases as repeat testers, like data element *Ever Had a Negative HIV Test (*#5), and is also used to calculate the inter-test interval between last negative and first positive tests for repeat testers.

### *Sources*

Sources for this variable may include patient interview, medical record review, physician note on a chart, ACRF, provider case report, field staff data collection form, databases at testing sites, PEMS or CTS forms, or other source. Field staff should receive training to note the date of last negative HIV test when following up on new cases for any reason.

### *Data Collection*

Because this is one of the most important data elements for incidence surveillance, extra effort should be made in collecting this date. Presence of a date of last negative HIV test classifies a person as a repeat tester, even if the patient's answer to question #5 is 'no'. During incidence estimation, the probability of being classified as recent by the BED assay is calculated separately for new testers and repeat testers. Any available information about testing history is useful in this calculation, including an approximate date or only the year tested, because it will classify a person as a repeat tester.

It is important to train interviewers, providers, and chart abstractors to record the last known date of a negative HIV test, even if the data collector does not know if negative results for the patient have occurred for later HIV tests performed at other facilities. If there are two dates for negative HIV tests, the most recent one should be entered. If the provider or chart abstractor is using the standard ACRF and has evidence of a documented negative HIV test with a test type, the date should be entered in the Laboratory Data section of the ACRF in the *Date of Last Documented Negative HIV Test* field. Otherwise, he or she should enter the date in the HIV Testing and ARV Use History section of the ACRF. Do not include dates of tests with unknown or indeterminate results.

### *Data Entry*

If the negative HIV test date is obtained from a documented laboratory report that contains test type (e.g., EIA), it should be entered in *Date of Last Documented Negative HIV Test* on the Lab Data tab of the ACRF form in eHARS. Any other dates of a negative HIV test which may come from patient self-report, provider report, chart notes, NHM&E/PEMS, or other databases should be entered on the TTH document, even if it is not known if it was absolutely the most recent test that the patient had. If another test date is found later, it should be entered on a separate eHARS document. If data entry person is unsure which form should be used to enter data, he or she should consult with the Incidence Surveillance Coordinator.

**Note: Until the implementation of eHARS version 4.0, it is recommended to also enter any documented negative laboratory test date found on the ACRF into the field for *Date of Last Negative HIV Test* on a TTH document, as well as on the Lab Data tab of the ACRF in the field for *Date of Last Documented Negative HIV Test*.** Because this data field is not properly contributing to the calculated variable for date of last negative test in eHARS, the date should also be entered on the eHARS TTH document until the defect is fixed in eHARS version 4.0. After version 4.0, the *Date of Last Documented Negative HIV Test* will be entered as any other negative HIV test. There will no longer be a separate field on the eHARS screen for this data element. There will be a calculated variable for the date of last negative HIV test before first positive test.

### *Examples by Source of Information:*

Patient Interview: Interviewer should record the date of the last known negative HIV test reported by the patient, even if the date only contains the year tested.

- Patient reported having a negative HIV test but cannot remember the date—interviewer should prompt patient to estimate the month and year tested. If unsuccessful, leave the date blank. Also record 'yes' for *Ever Had a Negative HIV Test* (#5).
- Patient recalled the date of one negative HIV test but was unsure if she had another one after that date—interviewer should record reported date.
- Interviewer found evidence of a previous negative test from another source, even if the patient forgot or was unaware of the test (e.g., didn't return for results)—interviewer should record the date. It is best practice to record the date on a separate TTH form. Retain patient's answer for *Ever Had a Negative Test* (#5).
- Patient reported she donated blood 3 months ago and says she must have been HIV negative or the blood bank would have called her—do not record this date. Do not assume that is evidence of a negative HIV test.

Medical Record Review: Chart abstractors should examine the chart for a provider note, patient self-report, or laboratory evidence of the last known negative HIV test.

- Provider note says 'patient reports negative HIV test in July 2007'—abstractor should record month and year of the reported test, not the date of the provider note.
- There is a report of a documented negative HIV test date—abstractor should record date in the Laboratory section of the ACRF along with the test type (e.g., HIV-1 EIA).
- A documented, negative HIV test is found in the medical record and recorded in the Laboratory section of the ACRF. A later patient-reported date is also described in the physician note. Record the later patient-reported date in the Testing and ARV Use History section of the ACRF or other data collection form.

Provider Report: Health care providers should be encouraged to ask patients about previous negative HIV tests and to record the most recent reported date, even if it only contains the year.

- Provider documented a laboratory result with test type—provider should report in the Laboratory section of the ACRF under *Date of Last Documented Negative HIV Test.*
- Provider from a clinic that runs a community HIV testing program consults a database of past tests and finds a date of last negative test for the patient but with no test type—provider should report the date in the Testing and ARV Use History section of the ACRF or other provider report form.
- Patient reported there is no previous HIV test—provider should leave date blank and report 'no' for *Ever Had a Negative HIV Test* (#5) on the provider report form.

NHM&E/PEMS: Enter the information obtained from NHM&E/PEMS (Part 3) or CTS form or database. If there is no information, leave blank.

If the value entered on any form or into any database is '999999' or there was no patient interview, no medical chart review, no provider report, and no other follow-up, the answer should be left blank.

## *1.7 TTH Data Element #7: Number of Negative HIV Tests within 24 months before First Positive HIV Test*

| **Variable name and format** | **UNUMTSTS (KNUMTSTS, UPNUMTSTS)** |
|---|---|
| | **IVR Values:** 1-99, **Refused** (R)**, Don't Know** (D)**;**<br>**eHARS Values:** 0-99, R, D |
| Wording on Pre-test form, IVR Database (UNUMTSTS): | 5b. For people who have NEVER had a positive test: In the past two years, how many times did you get tested for HIV? Today's test should be included for you in the count. |
| Wording on Pre-test form, IVR Database (UPNUMTSTS) | 5a. For people who have had a positive test: in the two years before your first positive HIV test (that is, the two years before the date in question 4b) how many times did you get tested for HIV? Your first positive test should be included for you in the count. |
| Wording on Post-test form, IVR Database (KNUMTSTS) | 9. In the two years before your first positive HIV test (in Question #3) how many times did you get tested for HIV? Your first positive HIV test (in Question #3) should be included for you in the count. |
| Wording on eHARS TTH | 7. Number of negative HIV tests within 24 months before first positive test (Do not include first positive HIV test) |
| Wording on ACRF | Number of negative HIV tests within 24 months before first positive test |
| Wording on PEMS Form | Number of tests in the two years before the current (or first positive) test. Include the current (or first positive) test. (Part 3) |
| | Values = 1-99 |
| Wording on proposed PEMS form | Number of negative HIV tests within 24 months before the current (or first positive) HIV test (Part 3) |
| | Values = 0-99, Don't know, Declined |

### *Description*

This variable, *Number of Negative HIV Tests within 24 Months before First Positive HIV Test*, is used to quantify the number of negative HIV tests within the 24 months preceding the first positive HIV test.

### *Purpose*

The purpose of this variable is to indicate testing frequency in the 24 months before the first positive HIV test, which is further used to calculate the inter-test interval for repeat testers in the absence of an actual date of last negative test. If a person reported having a previous test (*Ever had a Negative HIV Test = 'yes')* but *Number of Negative HIV Tests within 24 Months before First Positive HIV Test* is '0', then the inter-test interval is assumed to be >24 months. If there is no *Date of Last Negative HIV Test* and the value of the *Number of Negative HIV Tests within 24 Months before First Positive HIV Test* is 'R', 'D', or blank, multiple imputation is used to calculate the inter-test interval for repeat testers.

This variable is also used to classify cases as repeat testers when the *Number of Negative HIV Tests* indicates one or more previous negative tests and there is no information for *Ever had a Negative HIV Test (*#5) and *Date of Most Recent Negative Test (*#6).

### *History*

Historically the TTH forms and IVR database included the current positive HIV test in this count, which caused confusion. The value '1' was inconsistently reported and could not be used to classify individuals as a new tester or repeat tester. Starting on the date of conversion of HIS data to eHARS at a given surveillance area, only previous negative tests were counted in the newly-entered eHARS TTH documents; the current test, or first positive test, is no longer included. Field staff, providers, interviewers, and data entry staff need to be trained on the specific solution for the surveillance area, depending on the wording of data collection instruments.

For data collected but not entered prior to eHARS conversion, Incidence Surveillance Coordinators must correct the values for this data element before data are entered in eHARS or instruct data entry staff on how to enter the data correctly.

### *Sources*

Sources for this variable can be patient self-report, medical record review, doctor notes, and laboratory reports.

### *Data Collection*

Field staff who perform data collection and providers who complete case report forms need to be informed that as of the date the surveillance area converted incidence data to eHARS, this variable will no longer include the first positive HIV test.

The best source of this information is patient interview because a provider or medical record is unlikely to have the total of all tests in the 24-month period preceding the first positive HIV test. As a result this field is often blank or 'don't know'.

If the interviewer, abstractor or provider knows of at least one previous negative HIV test in the 24 months before first positive test, he or she should enter '1' even if there could be more unknown tests in this period. In the absence of a 'yes' for *Ever Had a Negative HIV Test* (#5) and a date for *Date of Last Negative HIV Test* (#6), entering a '1' for this variable classifies the person as a repeat tester. In addition, it indicates that a person received at least one negative HIV test in the two years prior to the HIV diagnosis date, which is used to estimate the inter-test interval for repeat testers.

If it is known that the patient did not have any HIV tests in the previous 24 months, then '0' should be recorded.

If the patient doesn't remember whether they had a negative test, or the provider or abstractor has no evidence about whether or not there was a previous test, enter 'don't know'. Do not record zero ('0'). Do not count an indeterminate test as a negative test. Ignore indeterminate tests.

**Note**: In analysis, when there is no known *Date of First Positive HIV Test* that is earlier than the *HIV Disease Diagnosis Date*, it is assumed that the *Date of First Positive HIV Test* is the same as the calculated eHARS *HIV Disease Diagnosis Date* (HIV_AIDS_DX_DT). If the patient/provider/chart reviewer does not know whether the patient ever had a previous positive HIV test, but he has information on the number of tests in the past two years, then this information should be captured. For example, the patient remembers one negative HIV test in the past 24 months but doesn't know if he had a previous positive. Entry of '1' allows us to categorize the patient as a repeat tester and estimate frequency of testing.

### *Data Entry*

Because data for this data element have been entered inconsistently, the guidance has changed, as of each surveillance site's date of conversion to eHARS, to stop including the first positive HIV test in the count. This change may require some data adjustment for this data element at the point of data entry for documents entered in eHARS on or after conversion to eHARS. The data entry person should consult with the Incidence Surveillance Coordinator for the appropriate data entry from the area's data collection forms.

**For areas that used the IVR Pre-test form:** At conversion, data for both of the Pre-test variables, UNUMTSTS and UPNUMTSTS, were converted to eHARS. Data for UNUMTSTS (number of tests for persons with no previous positive HIV test) appears at the top of the form and data for UPNUMTSTS (number of tests for persons with a previous positive test) appear under the Legacy Data and Other Optional Data Elements section at the bottom of screen. After conversion, do not enter any more data in the UPNUMTSTS field. Data can still be used for data analysis.

### *Examples by Source of Information:*

**Note**: These examples are based on the revised adult case report form and the eHARS TTH form, which do not include the first positive HIV test in this count.

Patient Interview:

- Patient never had an HIV test before—interviewer should record '0'.
- Patient indicated a previous positive but no previous negative HIV test—interviewer should record '0'.
- Patient self-reported previous testing—interviewer should record the number of negative tests that occurred in the 24 months before first positive HIV test.
- Patient reported having a negative test in the past but not in the last two years—question #5 should be 'yes' and number of tests should be '0'.
- Patient had one test in the 24 months before first positive but never got the results—do not count the test—interviewer should record 'don't know'.
- Patient had two negative tests and one HIV test for which she never got the results in the 24 months before first positive—interviewer should count the two negative tests before the current test but not the one with an unknown result.
- Patient knows he did not have a negative HIV test in the in the 24 months before first positive but doesn't know whether he was ever tested before then—interviewer should record '0' for number of negative tests and 'don't know' for *Ever Had a Negative HIV test.*
- Patient doesn't remember—interviewer should select 'don't know'.
- Interviewer finds evidence of one or more previous negative tests in the 24 month period—interviewer should record the number.
- Question was never asked—interviewer should leave blank.

Medical Record Review: Field staff needs to receive training that this variable changed as of the date of incidence conversion to eHARS and no longer includes the first positive HIV test in the count. It is generally difficult to ascertain the number of negative tests in the past two years from chart abstraction. In the absence of evidence to the contrary, use the current positive HIV test to calculate the number of negative tests in the 24 month interval before the positive test.

- There was evidence in the medical record of at least one previous negative test in the 24 months prior to the date of first positive HIV test (but the total number is unknown)—record '1'. Abstractor should try to find a *Date of Most Recent Negative Test* and enter for #6.
- There was evidence in the record of more than one negative test in the 24 month period—abstractor should record the number of tests.
- Clinic has a database with all previous tests conducted for the individual at that facility—abstractor should record the number of negative HIV tests in the 24 month period, even if there might have been other tests performed at other facilities.
- Chart abstractor does not know if the person had a previous positive HIV test, but finds evidence of two negative tests in the 24 months before the current positive HIV test—record '2'.
- Chart abstractor finds that there are no patient visits in the 24 months before first positive test—enter 'Don't Know'. Do not enter '0'.
- Patient reported to provider that he had no previous HIV tests other than an indeterminate HIV test a few weeks before the first confirmed positive HIV test—abstractor should record '0'. Ignore the indeterminate test.
- Patient had a false positive HIV test when she was pregnant 24 months ago and was confirmed to be HIV negative 6 months later. She had no other HIV tests until the current one. Record '1' for the negative test that occurred 18 months before the first positive test. Do NOT enter any information for the false positive test as it was not confirmed. Abstractors should ignore any false positive and indeterminate tests.
- There is no mention of the number of negative HIV tests in the record—abstractor should record 'don't know'.

Provider Report: Providers should be instructed that this variable changed as of the date of incidence conversion to eHARS to only include negative tests in the 24 months prior to the first positive HIV test.

- Provider knows the patient never had a previous test—report '0'.
- Patient self-reported at least one negative test in the 24 months preceding first positive test — report '1' (or more tests, depending on the number reported).
- Provider has no knowledge about previous negative HIV tests—record 'don't know'.
- Provider never asked the question of the patient—leave blank.

NHM&E/PEMS: Enter the information obtained from NHM&E/PEMS (Part 3) form or database. If there is no information, leave blank.

If there was no patient interview, no medical chart review, no provider report, and no other follow-up, the answer should be left blank.

### *1.8 TTH Data Element #8: Ever Taken Any Antiretroviral Medications (ARVs)?*

| **Variable name and format:** | **UHRT (KHRT) Yes/No/Refused/Don't Know** |
|---|---|
| Wording in IVR Database: | 7. In the past 6 months, have you taken any medicines shown in the picture on the last page to treat or try to prevent HIV or Hepatitis? |
| Wording on eHARS TTH | 8. Ever taken any antiretroviral medications (ARVs)? |

| | |
|---|---|
| | Values: 'N - No', 'Y - Yes', 'R – Refused', 'D – Do not know' |
| Wording on ACRF | Ever taken any antiretrovirals (ARVs)? |
| Wording on NHM&E/PEMS Form | Has client used or is client currently using antiretroviral medication (ARV)? (Part 3) |
| | Values: Yes, No, Don't know, Declined |

### *Description*

The variable, *Ever Taken Any Antiretroviral Medications*, is used to determine whether the patient took any antiretroviral medication to prevent or treat HIV or hepatitis at any time before the collection of the specimen used for the BED test. This variable is used in conjunction with the *Date ARVs First Began* and *Date of Last ARV Use* TTH variables and the STARHS specimen collection date.

A list of current medications used to treat HIV is available at:
http://www.crine.org/templates/cri/pdfs/cri_pillchart_jan09_ver3.pdf

A database to search medications by name is available at:
http://aidsinfo.nih.gov/DrugsNew/Default.aspx

### *Purpose*

This data element is important because ARV use may cause the BED results to appear 'recent' when the infection is not recent. It is critical to provide dates to know whether ARVs might have had any effect on the BED results. In the estimation model, for cases that have been exposed to ARVs in the 6 months prior to HIV disease diagnosis date, the BED results are set to missing and imputed. If ARVs were started after the collection of the specimen used for the BED assay or ended more than 6 months before that date, the BED results will be used as reported.

For Variant, Atypical, and Resistant HIV Surveillance (VARHS), this variable is used to determine eligibility. Any use of ARVs on or prior to the date of HIV diagnosis will make the case ineligible for VARHS.

Note that this data element is not being used to monitor treatment.

### *Sources*

Sources of the information may come from patient self-report, physician's notes, medical chart, pharmacy records, or the AIDS Drug Assistance Program (ADAP).

### *History*

This data element changed in 2007 when HIS transitioned from primarily conducting patient interviews to collecting data elements from all sources. At that time, it was decided to change the data element to '*Ever Taken Any ARVs*' instead of recording ARV use only in the six months prior to initial HIV diagnosis so that the variable could also be used for VARHS eligibility. This introduced some uncertainty with the timeframe of this variable. Now this data element covers any use of ARVs before or after diagnosis. However, only ARV use prior to the specimen collection date for the BED assay, which could include a specimen collected within 3 months after diagnosis, is used in calculations for incidence estimation.

### *Data Collection*

For any reported use of ARVs, even if ARV use began after the date of HIV diagnosis, record 'yes' but be sure to include *Date ARVs First Began* (#10) and *Date of Last ARV Use* (#11). Even if start/end dates are unknown, record 'yes'. When it is unknown whether the patient ever used ARVs, record 'don't know'. Do not assume that the absence of ARV use information indicates that the patient never used ARVs. **Note**: When there is no patient interview, it is difficult to know about previous ARV use.

If the patient has sporadic ARV use before HIV diagnosis, enter 'yes' and make an effort to obtain dates of first and last use. See #10 and #11 below. If the patient did not take ARVs until after HIV diagnosis, make an effort to obtain the month and year when medications began (even an estimated date) that is after the eHARS *HIV Disease Diagnosis Date*.

### *Data Entry*

ARV use is still relatively rare so this element is often blank or 'don't know'. Enter the data as recorded on the form. If there is a question, the data entry person should consult with the Incidence Surveillance Coordinator. If there is no information on the form, leave the field blank.

### *Examples by Source of Information:*

Patient Interview: If the patient answers 'yes', it is important to ask for approximate dates when ARVs began and ended, if appropriate (see #10 and #11).

- Patient reports taking prophylactic doses (prescribed or not) of ARVs—interviewer should enter 'yes'.
- ARV use occurred after the HIV diagnosis date but patient does not recall dates—interviewer should enter 'yes'. It is especially important for the interviewer to elicit approximate dates to show that ARV use occurred after HIV diagnosis date in this case. If no date is provided for *Date ARVs First Began*, then the specimen may not be eligible for VARHS.

Medical Record Review: Medical record abstractors should pay attention to dates of ARV use since many patients will be put on ARVs shortly after initial diagnosis.

- Abstractor finds ARV use before the first positive HIV test date—abstractor should enter 'yes'. Also record the dates *Date ARVs First Began* (#10) and *Date of Last ARV Use* (#11).
- Abstractor finds ARV use began shortly after diagnosis—abstractor should record 'yes' and include dates for #10 and #11.
- Chart note indicates patient was part of a Pre-exposure Prophylaxis (PrEP) clinical trial—abstractor should record 'yes' and include dates for #10 and #11.
- Chart note indicates the patient was being treated for hepatitis—check medications list to verify that the medication is an ARV and record 'yes' and include dates for #10 and #11. If the medication is not an ARV, record 'no'.
- Chart note indicates the patient never was on ARVs—enter 'no'.
- There is no mention of whether or not ARVs were used– enter 'don't know'.

Provider Report: Providers should be encouraged to ask patients if they ever used ARVs, including PrEP or non-prescribed ones, and report dates of ARV use.

- Provider knows patient has taken ARVs in the past but is uncertain of the timing—provider should record 'yes'.

- Provider started patient on ARVs after date of first positive HIV test—provider should record 'yes' and *Date ARVs First Began* (#10) and *Date of Last ARV Use* (#11).
- Patient was given ARVs after an occupational exposure—provider should report 'yes'.
- Patient reports taking trimethoprim-sulfamethoxazole (TMP-SMX, Bactrim, Septra, Co-trimoxazole), an antibiotic used to prevent *Pneumocystis carinii* Pneumonia (PCP) in HIV patients that is not an ARV—report 'no'.
- If the provider has no information about ARV use, enter 'don't know'.

NHM&E/PEMS: Enter the information obtained from NHM&E/PEMS (Part 3) or CTS form or database. If there is no information, leave blank.

If there was no patient interview, no medical chart review, no provider report, and no other follow-up, the answer should be left blank.

### *1.9 TTH Data Element #9: Name(s) of ARV Medication Taken*

| **Variable name and format:** | **UHRTA1 (MEDS) 2-digit code for various meds[1]** |
|---|---|
| Wording in IVR Database: | 7a. Which ones did you take? (If you are not sure of when you took the medicines, please include the ones you MIGHT have taken in the six months before your first positive test) |
| Wording on eHARS TTH | 9. If yes, name(s) of ARV medication taken |
| Wording on ACRF | If yes, ARV medications: |
| Wording on PEMS Form | If yes, specify antiretroviral medication (Part 3). **Note**: there is space for 4 medication codes which are listed on back of form |

[1]See Appendix B for list of medications in eHARS

#### *Description*

This variable, *Name(s) of ARV Medication Taken*, lists at least one of the ARV or hepatitis medications that the patient has taken but may not include all medications used.

#### *Purpose*

The purpose of this data element is to verify that at least one medication taken was actually an antiretroviral used to prevent or treat HIV or hepatitis. This is mostly used for verification during a patient interview. Chart abstractors and health care providers should be able to recognize the difference between ARVs and other medications, so this variable is less critical for collection. CDC has not used this variable for any data analysis, but individual surveillance areas may want to know all ARVs taken.

#### *Sources*

Sources of this information may be patient self-report, physician's notes, medical chart, or pharmacy records. A list of current medications used to treat HIV is available at:
http://www.crine.org/templates/cri/pdfs/cri_pillchart_jan09_ver3.pdf
and a database to search names of medications is available at
http://aidsinfo.nih.gov/DrugsNew/Default.aspx.

### *Data Collection*

It is not necessary to list every drug combination that may have been used; record at least one ARV medication. It is important to record dates of first and last use (#10 and #11).

### *Data Entry*

Enter the medications as recorded on the form. In eHARS, hold down the Control key and select all of the recorded medications from the pick list. If the ARV does not appear in the pick list, check the Web sites listed above to verify the medication is an ARV. If the ARV is a new ARV drug that does not appear on the eHARS list, enter 'other'. If there is a question whether the medication is an ARV, the data entry person should consult with the Incidence Surveillance Coordinator. If there is no information on the form, leave the field blank.

### *Examples by Source of Information:*

Patient Interview: For patients that report 'yes' to *Ever Taken Any Antiretroviral Medications (*#8), ask which ARVs they have taken. If patient is uncertain of the names, show a pill chart. Only one ARV medication name is required. If patient reports taking only trimethoprim-sulfamethoxazole (TMP-SMX), also known as Bactrim, Septra, or Co-trimoxazole, an antibiotic used to prevent *Pneumocystis carinii* Pneumonia (PCP) in persons living with HIV and which is not an ARV—interviewer should report 'no' to *Ever Taken Any Antiretroviral Medications* (#9) and leave this field blank.

- Patient names medications that are not used to treat or prevent HIV or hepatitis, probe further. If it is apparent that the person did not take ARVs, interviewer should report 'no' for #8 and leave this variable blank.
- Patient does not know the names of medications taken—select 'unspecified' or leave blank if it is suspected that the patient was not taking ARVs. Incidence Surveillance Coordinators need to consider any notes made by the interviewer and exercise judgment in determining whether the answer to *Ever Taken Any Antiretroviral Medications (*# 8) should be 'yes' or 'don't know'.

Medical Record Review: Examine physician notes (including patient self-report), noted prescriptions, and pharmacy records to see if patient was taking an antiretroviral medication. List at least one med taken. Record dates of use.

- Provider note indicates patient was on ARVs before seeking treatment—abstractor should list at least one medication.
- Provider note indicates patient has taken non-prescribed drugs (e.g., medications obtained outside of a clinical setting or "on the street"), for prevention of HIV—abstractor should list medication or, if unknown, record 'unspecified'.

Provider Report: Instruct providers that they do not have to list all drugs, that one is sufficient.

- Patient has moved to this state and has been on various ARVs for over 5 years—provider should enter at least one of the medications.
- Patient has taken ARVs in the past but provider is not sure which ones—provider should record 'unspecified'.

NHM&E/PEMS: Enter the information obtained from NHM&E/PEMS (Part 3) or CTS form or database. If there is no information, leave blank.

If there was no patient interview, no medical chart review, no provider report, and no other follow-up, the answer should be left blank.

## *1.10 TTH Data Element #10: Date ARVs First Began*

| **Variable name and format:** | **UHRTBD (KHRTBD) mm/dd/yyyy** |
|---|---|
| Wording in IVR Database: | 7b. What was the first day you took any of the medications shown in the pictures? Please make your best guess if you are not sure. |
| Wording on eHARS TTH | 10. Dates ARVs taken: Date first began |
| Wording on ACRF | Dates ARVs taken: Date first began |
| Wording on NHM&E/PEMS Form | Date ARV began |

### *Description*

This date, *Date ARVs First Began*, represents the earliest date of any ARV use.

### *Purpose*

The purpose of this data element is to determine whether the patient took any antiretroviral medication to prevent or treat HIV or hepatitis before or after the date that the specimen used for the BED test was collected. This variable is important because ARV use may cause the BED results to appear 'recent' when the infection is not recent. Therefore, in incidence estimation, for cases that have been exposed to ARVs in the 6 months prior to BED specimen collection date, the BED results are set to missing and imputed. For ARV start dates after the specimen collection date, the reported BED results are maintained.

This date is also used to determine if the patient was on ARVs on or prior to the date of diagnosis in order to assess eligibility for VARHS.

### *Sources*

Source for this item includes: patient self-report, chart abstraction, doctor notes, ADAP database, and pharmacy records.

### *Data Collection*

This data element is important for determining the period when any ARV use started. Medical record abstractors should pay attention to start dates of ARV use, even those after initial HIV diagnosis. Providers should be informed that the start date is the critical piece of information. Record the earliest month and year of ARV use. Note that this date is not necessarily related to the time that the specific medication named in #9 was taken.

Do not be concerned if ARV use has been sporadic. It is true that some BED results that were not affected by ARVs may not be used because the *Date ARVs First Began* was well before HIV diagnosis and the *Date of Last ARV Use* from a subsequent use is after diagnosis, but this relatively rare scenario is an acceptable limitation.

### *Data Entry*

Enter the data as recorded on the form. If there is a question, the data entry person should consult with the Incidence Surveillance Coordinator. If there is no information on the form, leave the field blank.

### *Examples by Source of Information:*

Patient Interview: Enter the first date the patient used ARVs. Ask about prophylactic use (prescribed or not).

- Patient used non-prescribed ARVs intermittently in the past year—interviewer should record the first month/year of ARV use.
- Patient is uncertain about the start date—interviewer should use prompts to obtain the estimated month and year.
- Patient was diagnosed in another state 2 years ago and took ARVs until he moved to your jurisdiction 9 months ago. Now he has been retested in order to enter care and the provider has put the patient back on ARVs. Provider should record the earliest known date of ARV use.
- Patient cannot remember date—leave blank.

Medical Record Review: Examine physician notes (including patient self-report), noted prescriptions, and pharmacy records to find dates when patient was taking any antiretroviral medication. Record the first date ARV use began before HIV diagnosis.

- ARV use began before the eHARS HIV diagnosis date and continued after diagnosis—abstractor should record the earliest date.
- ARV use started later than HIV diagnosis—abstractor should record the earliest date.
- ARV use was intermittent for several years—abstractor should record the earliest date.
- ARVs used but dates are unavailable—leave date blank.

Provider Report: Providers should be encouraged to ask for dates of ARV use for patients that report ever using ARVs. Providers should be reminded that this date is not related to the time that the specific medication named in #9 was taken.

- Provider should record first known date of ARV use, even if use has been sporadic.
- If the provider has no information about ARV use dates—leave date blank.

NHM&E/PEMS: Enter the information obtained from NHM&E/PEMS (Part 3) or CTS form or database. If there is no information, leave blank.

If the value is '999999' or there was no patient interview, no medical chart review, no provider report, and no other follow-up, the answer should be left blank.

## *1.11 TTH Data Element #11: Date of Last ARV Use*

| **Variable name and format:** | **UHRTED (KHRTED) mm/dd/yyyy** |
|---|---|
| Wording in IVR Database: | 7b. What was the last day you took any of the medications shown in the pictures? Please make your best guess if you are not sure. |
| Wording on eHARS TTH | 11. Date of last ARV use |
| Wording on ACRF | Date of last use |
| Wording on NHM&E/PEMS Form | Date of last ARV use (Part 3) |

### *Description*

This variable, *Date of Last ARV Use*, represents the date when ARVs were last taken by the patient. The *Date Patient Reported Information* should be recorded as the *Date of Last ARV Use* if the patient was still on ARVs as of that date.

### *Purpose*

The purpose of this data element is to determine whether the patient took any antiretroviral medication to prevent or treat HIV or hepatitis in the 6 months before the date that the specimen used for the BED test was collected. This variable is important because ARV use may cause the BED results to appear 'recent' when the infection is not recent. Therefore, for incidence estimation, for cases that have been exposed to ARVs in the 6 months prior to BED specimen collection date, the BED results are set to missing and imputed in the incidence estimation program.

### *Sources*

Sources for this item include: patient self-report, chart abstraction, doctor notes, ADAP database, and pharmacy records.

### *Data Collection*

Record the last date when the patient was known to be taking ARV medications, prescribed or not. If ARVs are currently being taken, record the date when the patient was last known to be taking ARVs. That is likely to be the same as *Date Patient Reported Information (#2),* the date of the interview, chart note or provider encounter.

### *Data Entry*

Enter the data as recorded on the form. If there is a question, the data entry person should consult with the Incidence Surveillance Coordinator. If there is no information on the form, leave the field blank.

### *Examples by Source of Information:*

Patient Interview: Enter the last date the patient was known to use ARVs. If the patient is currently on ARVs, that date would be the interview date.

- Patient had intermittent use—interviewer should select the last date used.
- Patient took some ARVs to prevent infection but stopped two years ago—interviewer should record the approximate date of last use, even if it is only the year of use.
- Patient is uncertain about the last ARV use date—interviewer should use prompts to obtain the estimated month and year.
- Patient still taking ARVs—interviewer should record date of interview.

Medical Record Review: Examine physician notes (including patient self-report), noted prescriptions, and pharmacy records to find date when patient was last known to be taking ARVs or stopped taking antiretroviral medication.

- Patient still taking ARVs—record the latest date of physician note or last prescription.
- Patient was prescribed ARVs but discontinued use—abstractor should record the date of last known use.

Provider Report: For patients that ever used ARVs, providers should be encouraged to ask for the last date of ARV use.

- Provider has no information about ARV use dates—leave blank.
- Patient is still taking ARVs—report date of last known use or date of encounter.

NHM&E/PEMS: Enter the information obtained from NHM&E/PEMS (Part 3) or CTS form or database. If there is no information, leave blank.

If the value is '999999' or there was no patient interview, no medical chart review, no provider report, and no other follow-up, the answer should be left blank.

**For required STARHS specimen and laboratory results data elements, skip to page 44**.

# 2. Legacy Data and Other Optional TTH Data Elements

## *2.1 TTH Data Element #12: Are You Now Taking Any ARVs?*

| **Variable name and format:** | **QHRTNW (KHRTNW) Yes/No/Refused/Don't Know** |
|---|---|
| Wording in IVR Database[1]: | 7c. Are you now taking any of the medications shown in the pictures? |
| Wording on eHARS TTH: | 12. Are you now taking any ARVs? |
| | Values: 'N – No', 'Y – Yes', 'R – Refused', 'D – Do not know' |
| Wording on ACRF | Not in ACRF TTH section |

[1] From Pre-test version in IVR database

### *Collection and Data Entry/History*

The purpose of the *Are You Now Taking Any ARVs* variable was to ascertain if the patient was on ARVs at the time the specimen was collected for the BED test because ARV use could cause false 'recent' results of the BED test. This variable became optional as of 2007. Since TTH information could be obtained through chart review, provider report, or interview before or after the HIV diagnosis, this variable lacked a temporal component. For example, if the information was obtained through medical record abstraction 4 months after diagnosis, the patient could be 'currently' on ARVs, though they were not taking them at the time the specimen was collected for diagnosis and BED testing. The dates specified in the *Date ARVs First Began* and *Date of Last ARV Use* variables are more precise indicators.

## *2.2 TTH Data Element #13: Ever Tested for HIV Before Today? (Legacy Pre-test form)*

| **Variable name and format:** | **(UPTESTS) Yes/No/Refused/Don't Know** |
|---|---|
| Wording in IVR Database: | 4. Have you ever been tested for HIV before today? |
| Wording on eHARS TTH: | 13. Ever been tested for HIV before today? |
| | Values: 'N – No', 'Y – Yes', 'R – Refused', 'D – Do not know' |
| Wording on ACRF | Not in ACRF TTH section |

### *Collection and Data Entry*

The *Ever Tested for HIV before Today* variable appeared on the pre-test form and was not used by the majority of incidence surveillance areas. It was part of a series of questions for a patient interview. When it was used as part of an interview, the variable was part of a skip pattern for a series of questions. For analysis, if a patient reported 'No', then he or she should be classified as a new tester without a previous positive test.

### *History*

*Ever Tested for HIV before Today* became optional in 2007 because better information could be obtained from other variables regarding first positive and first negative tests.

## *2.3 TTH Data Element #14: Date of First HIV Test Ever*

| **Variable name and format:** | **UFTSTD (KFTSTD) mm/dd/yyyy** |
|---|---|
| Wording in IVR Database: | 6. When was the very first time you ever got tested for HIV (when you got the test, not when you got the results)? Please make your best guess if you are not sure. |
| Wording in eHARS: | 14. When was the first time you ever got tested for HIV? |
| Wording on ACRF | Not in ACRF TTH section |

### *Collection and Data Entry/History*

The purpose of the *Date of First HIV Test Ever* variable was to ascertain previous testing. However, it lacked the specificity of whether the test was positive or negative. This data element became optional in 2007 because better information could be obtained from other variables regarding date of first positive and first negative HIV tests.

## *2.4 TTH Data Element #15: Was the First Positive HIV Test Anonymous?*

| **Variable name and format:** | **UFPOSA (KFPOSA) Yes/No/Refused/Don't Know** |
|---|---|
| Wording in IVR Database: | 4c. When you first tested positive for HIV were you given a number or code to use to get your results instead of your name? |
| Wording on eHARS TTH: | 15. When you first tested positive for HIV, was the HIV test an anonymous test? |
| Wording on ACRF | Not in ACRF TTH section |

### *Collection and Data Entry/History*

The purpose of the *Was the First Positive HIV Test Anonymous* variable was to clarify the reason there was a positive HIV test before the HIV diagnosis date. There are a number of reasons that a previous positive test may not have been reported before; for example, the person was tested in another state, was tested in a state that did not have HIV reporting, or had an anonymous test. This data element became optional in 2007.

## *2.5 TTH Data Element #16: Number of Tests 2 years before Previous Positive HIV Test (Legacy Pre-test form)*

| **Variable name and format:** | **UPNUMTSTS** |
|---|---|
| Wording in IVR Database[1]: | 5a. For people who have had a positive test: In the two years before your first positive HIV test (that is, the two years before the date in question 4b), how many times did you get tested for HIV? Your first positive test should be included for you in the count. |
| Wording on eHARS TTH: | 16. For persons who had a previous positive test (Pre-test - Legacy only; enter new data in #7): In the two years before first positive test, how many times did you get tested for HIV? |
| | Values: 1-99; R, D |
| Wording on ACRF | Not in ACRF TTH section |

[1] From Pre-test version in IVR database

### *Collection and Data Entry/History*

The *Number of Tests 2 Years before Previous Positive HIV Test* variable appeared on the Pre-test form and was not used by the majority of incidence surveillance areas. The Pre-test form had two separate variables. This variable (UPNUMTSTS) was for persons with a previous positive test where the first positive test would be included in the count. The other variable (UNUMTSTS) was for persons with no previous positive tests who were interviewed before the HIV test and did not know they were positive. In the latter case, the first positive test had to be added to the count later. The purpose of this variable is the same as *Number of Negative HIV Tests within 24 Months before First Positive HIV Test* (#7).

**Note**: For areas that used the Pre-test form, during eHARS conversion data for this variable were converted to the variable UPNUMTSTS (eHARS column name) which can be used in analyses for data entered prior to conversion. After conversion to eHARS, do not enter data to this field. All counts of previous negative tests are now entered in the required variable, *Number of Negative HIV Tests within 24 Months before First Positive HIV Test* (#7).

## *2.6 TTH Data Element #17: Reason for Today's HIV Test (Legacy Pre-test form)*

| **Variable names and format:** | **UREAS3_ 1 through UREAS3_5, UR3_5SP Yes/No/Refused/Don't Know** |
|---|---|
| Wording in IVR Database: | 3. Why are you getting the HIV test today? Are you getting the test… |
| Wording in eHARS TTH | 17. Reason for getting today's HIV test (Legacy Pre-test form): |
| Wording on ACRF | Not in ACRF TTH section |

### *Collection and Data Entry*

The purpose of these variables, collectively named *Reason for Today's HIV Test,* was to ascertain whether the patient had reasons for getting an HIV test that would affect the probability of being tested in the BED window period, such as being concerned about a possible recent exposure. These variables were collected for reason for testing at the time of the interview, which occurred before the test results were known.

There are 5 standard yes/no questions and 19 other reasons for testing. The following are the standard reasons, using current eHARS wording, with the corresponding variable in parentheses:

a. Think you might have been exposed to HIV in the past 6 months? (UREAS3_1)
b. Get tested on a regular basis and it is time to get tested again? (UREAS3_2)
c. Just checking to make sure you are HIV negative? (UREAS3_3)
d. Required to get the test by insurance, military, court, or other agency? (UREAS3_4)
e. Other reason you wanted to get tested? (UREAS3_5)
f. If other reason, describe: _____ (UR3_5SP)

The variable for *#17f., Other Reason for Today's Test (describe),* is an open text field in eHARS. However, in the IVR database, the following were offered as commonly used other reasons. It is recommended to use these (entered exactly as written or beginning the same way) because they can be analyzed more easily.

The other reasons are:

1. MD recommendation / rule out HIV diagnosis
2. Patient has symptoms, recent illness, wt loss, OI
3. Current STD and/or STD screening
4. Named as contact or partner/ ex is HIV+/symptomatic
5. Possible exposure >6 months ago
6. Incarceration
7. Blood/plasma donation or referred by blood bank
8. Prenatal screening or pregnancy
9. Hospitalization, pre-op test, other procedure
10. Entry to drug/alcohol treatment
11. Requested by partner, family, other
12. Confirm previous HIV + test
13. Needed to initiate care
14. Starting/ending a relationship
15. HIV vaccine trial
16. Immigration screen
17. Community screening /free test /test offered
18. Occupational exposure
19. NAAT testing

### *History*

In developing the first incidence estimates, CDC found that reasons for testing, such as possible exposure, were not associated with a higher likelihood of having a BED recent result. As a result, these data elements were no longer required as of 2007, but some surveillance areas continue to collect this information for their own purposes.

These particular variables, *Reason for Today's HIV Test,* only appeared on the Pre-test form. For areas that used the Pre-test form, these reasons for testing were for newly diagnosed persons that did not self-report a previous positive HIV test. For persons that reported having a previous positive test, a second question about reasons was included on the Pre-test form, which appears under the following data element, *Reason for First Positive HIV Test* (#18).

## *2.7 TTH Data Element #18: Reason for First Positive HIV Test*

| **Variable name and format:** | **URS4E_1 through URS4E_5 (KREAS6_1 thru _5) Yes/No/Refused/Don't Know** |
|---|---|
| Wording in IVR Database: | 4e. Why you got your first positive test? Did you get the test… |
| Proposed wording in eHARS | 18. Reason for getting the first positive HIV test |
| Wording on ACRF | Not in ACRF TTH section |

### *Collection and Data Entry*

The purpose of these variables, collectively named *Reason for First Positive HIV Test*, was to ascertain whether the patient was motivated to get tested by an exposure or had other reasons for testing that would affect the probability of being tested in the BED window period. The above named variables were collected for reason for testing at the time of the self-reported first positive test. There are 5 standard reasons (using current eHARS wording), with each a separate variable named in parentheses:

a. Thought that you might have been exposed to HIV in the 6 months before the test? (URS4E_1)
b. Got tested on a regular basis and it was time to get tested again? (URS4E_2)
c. Just checking to make sure you were HIV negative? (URS4E_3)
d. Required to get the test by insurance, military, court, or other agency? (URS4E_4)
e. Other reason you wanted to get tested? (URS4E_5)
f. If other reason, describe: _____ (URS4E_5SP)

The variable for *18f., Other Reason for First Positive HIV Test (describe)*, is an open text field in eHARS. However, in the IVR database, the following were offered as commonly used other reasons. It is recommended to use these (entered exactly as written or beginning the same way) because they can be analyzed more easily. The other reasons are:

1. MD recommendation / rule out HIV diagnosis
2. Patient has symptoms, recent illness, wt loss, OI
3. Current STD and/or STD screening
4. Named as contact or partner/ ex is HIV+/symptomatic
5. Possible exposure >6 months ago
6. Incarceration
7. Blood/plasma donation or referred by blood bank
8. Prenatal screening or pregnancy
9. Hospitalization, pre-op test, other procedure
10. Entry to drug/alcohol treatment
11. Requested by partner, family, other
12. Confirm previous HIV + test
13. Needed to initiate care
14. Starting/ending a relationship
15. HIV vaccine trial
16. Immigration screen
17. Community screening /free test /test offered
18. Occupational exposure
19. NAAT testing

### *History*

In developing the first incidence estimates, CDC found that reasons for testing, such as exposure, were not associated with a higher likelihood of having a BED recent result. As a result, these data elements were no longer required as of 2007, but some surveillance areas continue to collect this information.

The questions were asked once on the Post-test form but were included on the Pre-test form in addition to the reasons for "today's test" for those persons who reported having a previous positive HIV test. At conversion, the Post-test variables for *Reason for First Positive HIV Test* (KREAS6_1 through _5)

were mapped to the Pre-test variables for *Reason for First Positive HIV Test* (URS4E_1 through URS4E_5).

## *2.8 TTH Data Element #19: Name of Facility Where First Tested Positive for HIV*

| **Variable name and format:** | **UFPS_SITE (KFPS_SITE) Text field** |
|---|---|
| Wording in IVR Database: | 4d. What was the name of the place where you got your first positive HIV test (on the date in question 4b)? For example, this could be the name of a health clinic, blood bank, doctor's office, or STD clinic. Site Name: |
| Wording in eHARS TTH: | 19. Name of facility where first tested positive for HIV |
| Wording on ACRF | Not in ACRF TTH section |

### *Collection and Data Entry/History*

The *Name of Facility Where First Tested Positive for HIV* variable is for local use to record the facility where the person first tested positive for HIV. It can be used to monitor the sources of new diagnoses as well as provide information for follow-up activities (e.g., chart review). This data element is not transferred to CDC through eHARS.

This data element was no longer required as of 2007, but some surveillance areas continue to collect this information for local use.

## *2.9 TTH Data Element #20: State of Facility Where First Tested Positive for HIV*

| **Variable name and format:** | **UFPS_STATE (KFPS_STATE) 2-digit State code** |
|---|---|
| Wording in IVR Database: | 4d. Site State: |
| Wording in eHARS TTH: | 20. State of facility where first tested positive for HIV |
| Wording on ACRF | Not in ACRF TTH section |

### *Collection and Data Entry/History*

The *State of Facility Where First Tested Positive for HIV* variable is used to record the state where the person first tested positive for HIV. It can be used to explain why the self-reported diagnosis date is different from the eHARS *HIV Disease Diagnosis Date*. This data element is not transferred to CDC.

This data element is no longer required as of 2007, but some surveillance areas continue to collect this information for local use. For persons previously tested in other countries, "Foreign Country" is not an available choice on the eHARS drop down menu, but information can be stored in a local field.

## *2.10 TTH Data Element #21: Type of Facility Where First Tested Positive for HIV*

| **Variable name and format:** | **UFPSTYP (KFPS) eHARS FACILITY_TYPE code** |
|---|---|
| Wording in IVR Database: | 4d. Site Type: |
| Wording in eHARS TTH: | 21. Type of facility where first tested positive for HIV |
| Wording on ACRF | Not in ACRF TTH section |

### *Collection and Data Entry/History*

The *Type of Facility Where First Tested Positive* variable represents the type of testing facility where the person was first tested for HIV. This data element was to be used, in conjunction with the eHARS *Facility Type at HIV Diagnosis*, as a variable used in multiple imputation of missing information for testing group and BED results during calculation of incidence estimates, but because the variable is also included on the ACRF in eHARS, it is no longer required for HIS as of 2007. Some surveillance areas continue to collect this information for local use.

## *2.11 TTH Data Element #22: Name of Facility Where Last Tested Negative for HIV*

| **Variable name and format:** | **ULSTNGS_SITE (KLSTNGS_SITE) text** |
|---|---|
| Wording in IVR Database: | 4g. What was the name of the place where you had your last negative HIV test? For example, this could be the name of a health clinic, blood bank, doctor's office, or STD clinic. Site Name: |
| Wording in eHARS TTH: | 22. Name of facility where last tested negative for HIV |
| Wording on ACRF | Not in ACRF TTH section |

### *Collection and Data Entry*

The *Name of Facility Where Last Tested Negative for HIV* variable represents the place where the person last had a negative HIV test. This information is not transferred to CDC through eHARS. This data element is no longer required as of 2007, but some surveillance areas continue to collect this information for local use.

## *2.12 TTH Data Element #23: State of Facility Where Last Tested Negative for HIV*

| **Variable name and format:** | **ULSTNGS_STATE (KLSTNGS_STATE) 2-digit State code** |
|---|---|
| Wording in IVR Database: | 4g. Site State: |
| Wording in eHARS TTH: | 23. State of facility where last tested negative for HIV |
| Wording on ACRF | Not in ACRF TTH section |

The *State of Facility Where Last Tested Negative for HIV* variable represents the state where the person last had a negative HIV test. This data element is no longer required as of 2007, but some surveillance areas continue to collect this information for local use. For persons previously tested in other countries, "Foreign Country" is not an available choice on the eHARS drop down menu, but information can be stored in a local field.

### *2.13 TTH Data Element #24: Type of Facility Where First Tested Negative for HIV*

| **Variable name and format:** | **ULSTNGS (KLSTNGS) eHARS FACILITY_TYPE code** |
|---|---|
| Wording in IVR Database: | 4g. Site Type: |
| Wording in eHARS TTH: | 24. Type of facility where last tested negative for HIV |
| Wording on ACRF | Not in ACRF TTH section |

The *Type of Facility Where Last Tested Negative for HIV* variable represents the type of facility where the person last had a negative HIV test. This data element is no longer required as of 2007, but some surveillance areas continue to collect this information for local use.

## 3. IVR Database TTH Variables Not Included in eHARS TTH Document

### *3.1 Date of HIV Test (reference date)*

| **Variable name and format:** | **UPTESTD (KPTSTD) mm/dd/yyyy** |
|---|---|
| Wording in IVR Database: | 2. Date of this HIV test |
| Wording in eHARS TTH: | Not in eHARS |
| Wording on ACRF | Not in ACRF TTH section |

#### *Collection and Data Entry/History*

The purpose of the *Date of HIV Test* variable was to assist interviewers and field staff in the collection of data related to TTH around time of HIV diagnosis. It was used to note the date of the current positive HIV test in order to ask questions related to events before this date during an interview or chart abstraction. This variable became optional in 2007 and does not appear in eHARS. Data will not be converted into eHARS.

# 4. Required and Optional Data Elements for STARHS Specimen Information and Results

These data elements are presented in the order that they appear on the Lab Data tab of the eHARS Laboratory Report document. In eHARS, STARHS specimen and results information are entered like any other laboratory test. Therefore, some of these variables have multiple uses, depending on the type of laboratory test that is selected, and will have generic labels on the eHARS Laboratory document. When a STARHS laboratory test type is selected in eHARS, it opens a screen with all of the STARHS laboratory variables. Certain variables are required to enter a laboratory document and are noted here as 'Required for eHARS.' Variables that are required for HIS are so designated.

There can be multiple laboratory documents for the same specimen; for example, one for specimen information entered on one date and one for the results imported at a later date, each with the same STARHS ID. For analysis, the data from all Lab documents for a unique STARHS ID will be combined.

## *4.1 STARHS Laboratory Name*

| **Variable name and format:** | **CLIA_UID[1] (LABID)[2] CLIA code for STARHS lab** |
|---|---|
| Wording in IVR Database: | LabID |
| | IVR Values:<br>21D0649758 – Maryland Department of Health Laboratory<br>50D0661430 – University of Washington Laboratory<br>33D0654341 – Wadsworth Center–Biggs Laboratory (Legacy)<br>eHARS Value:33D2005937 – Wadsworth Center–David Axelrod Institute (Use in eHARS) |
| Wording in eHARS Lab Document: | Lab Name |

[1] eHARS variable

[2] IVR database variable

### *Description*

For HIV incidence surveillance, the *STARHS Laboratory Name* variable represents the name of the laboratory that conducted STARHS testing. Currently, all tests are conducted at the single CDC-funded STARHS laboratory using BED HIV-1 Capture EIA (BED), but in the past some other laboratories performed STARHS testing using Vironostika HIV-1 EIA (Vironostika-LS).

The list of laboratory names in eHARS is created by the local eHARS Administrator. For most STARHS specimens, the value for *STARHS Laboratory Name* will be the name of the CDC-funded STARHS laboratory, currently the Wadsworth Center – David Axelrod Institute.

### *Purpose*

The purpose of this variable is to identify the name of the laboratory which performed STARHS testing. In the IVR and eHARS databases, the laboratory name is stored as a unique Clinical Laboratory Improvement Amendments (CLIA) code assigned to that laboratory.

### *Source*

Sources for *STARHS Laboratory Name* are tracking databases and the STARHS results file received from the STARHS laboratory.

### *Data Entry*

The *STARHS Laboratory Name* is the name of the laboratory performing the HIV-related test, in this case, the STARHS testing algorithm. This data element may be manually entered or imported. The *STARHS Laboratory Name* can be selected from the list of laboratory names that appear on the eHARS screen (user interface) in the local eHARS installation. The value can also be imported from the STARHS Results file received from the STARHS laboratory using the All Document Import (ADI) Laboratory Report Default template. The SAS program used to import STARHS data will translate the values to the appropriate CLIA code.

### *Examples*

- Wadsworth Center – David Axelrod Institute – 33D2005937

## *4.2 Source Lab Specimen ID*

| **Variable name and format:** | **SAMPLE_ ID (LSRCEID)** |
|---|---|
| Wording in IVR Database: | Source Lab Specimen ID |
| Wording in eHARS Lab Document: | Sample ID (Specimen) |

### *Description*

The *Source Lab Specimen ID* variable was originally intended to capture the specimen ID from the originating laboratory for a STARHS specimen, usually a commercial laboratory. However, this data element is for local use and can be used in any way that facilitates identifying a specific specimen. Specimens may have several specimen numbers.

### *Purpose*

This variable is one of two used locally to track specimen identifiers: *Source Lab Specimen ID* (LSRCEID) for the originating or commercial laboratory and *State Lab Specimen ID* (SSTATEID) for public health laboratory identifiers. In eHARS, there are two generic specimen IDs that can be used, *Sample ID* (SAMPLE_ID) and *Accession Number* (ACCESSION_NO). At conversion to eHARS, the IVR variable, *Source Lab Specimen ID*, will be mapped to SAMPLE_ID and *State Lab Specimen ID* will be mapped to ACCESSION_NO. CDC does not use these data elements.

### *Data Entry*

Local surveillance areas should decide which specimen ID will be entered for this variable. The main goal is to be consistent; therefore, these decisions should be included in the Standard Operating Procedures (SOPs). These data may be manually entered or imported using the ADI Laboratory Report Default template.

### *4.3 State Lab Specimen ID (or Other Specimen ID)*

| **Variable name and format:** | **ACCESSION_NUMBER (SSTATEID)** |
|---|---|
| Wording in IVR Database: | State Lab Specimen ID |
| Wording in eHARS Lab Document: | Accession Number |

#### *Description*

The *State Lab Specimen ID* variable was intended to capture the specimen ID from the State Public Health laboratory. However, this data element is for local use and can be used in any way that facilitates identifying a specific specimen.

#### *Purpose*

This variable is also used to track specimens, particularly from State Public Health laboratories. At conversion the IVR variable, SSTATEID, will be mapped to the eHARS field labeled Accession Number (ACCESSION_NO).

#### *Data Entry*

Local surveillance areas should decide what specimen IDs will be entered for this variable. To be consistent, these decisions should be included in the SOPs. These data may be manually entered or imported using the ADI Laboratory Report Default template.

### *4.4 Date of Specimen Collection (Required for HIS)*

| **Variable name and format:** | **SAMPLE_DT[1] (LDTEOBT)[2] mm/dd/yyyy** |
|---|---|
| Wording in IVR Database: | Date of Specimen Collection |
| Wording in eHARS Lab Document: | Collection date |

#### *Description*

The *Date of Specimen Collection* variable is the date the specimen was drawn for the HIV-related test that was used for STARHS testing. This is the draw date of the remnant specimen from the initial HIV diagnostic specimen or another HIV-related test within 3 months of HIV diagnosis.

#### *Purpose*

The purpose of this data element is to determine eligibility of the specimen for STARHS testing. It is an HIS-required variable for STARHS specimens. This data element is compared to the eHARS calculated *HIV Disease Diagnosis Date* (Calc OBS 285) which includes earliest AIDS date in the calculation. If the date is later than 3 months from HIV diagnosis, this specimen should not be sent for STARHS testing. If it is tested, the STARHS result from this specimen is not used in the HIV incidence estimation.

#### *Sources*

Sources of this date are paper or electronic laboratory reports and laboratory tracking databases.

### *Data Entry*

The *Date of Specimen Collection* variable can be entered manually or imported using the ADI STARHS Lab template. This variable is required for HIS. When entering any lab document in eHARS, if there is no specimen collection date entered there will be a warning given after submitting the document. However, the specimen data will be stored in eHARS. For HIS, there must be a collection date on at least one lab document for a given STARHSID for the results to be used in data analysis or incidence estimation.

## *4.5 STARHS Result Date*

| **Variable name and format:** | **RESULT_RPT_DT (TESTDATE) mm/dd/yyyy** |
|---|---|
| Wording in IVR Database: | Test Date |
| Wording in eHARS Lab Document: | Result Date |

### *Description*

The *STARHS Result Date* variable is the date that the STARHS lab performed the STARHS test on the remnant sample.

### *Purpose*

This variable is not used for data analysis but is useful for data timeliness evaluation, monitoring the time from reporting date or shipping date to STARHS testing.

### *Sources*

The source of this data element is the STARHS Results file sent by the STARHS laboratory.

### *Data Entry*

These data may be manually entered in eHARS or imported using ADI Laboratory Report Default template.

## *4.6 Received Date*

| **Variable name and format:** | **RECEIVE_DT mm/dd/yyyy** |
|---|---|
| Wording in IVR Database: | Not in IVR Database |
| Wording in eHARS Lab Document: | Received Date |

### *Description*

The *Received Date* variable is an eHARS data element for the date the laboratory result was received at the health department. This variable did not appear in the IVR database.

### *Purpose*

This data element is not used by incidence surveillance. Local surveillance areas may choose to use it to monitor the time when STARHS results were received from the STARHS lab. Otherwise, the variable can be left blank.

### *Data Entry*

These data may be manually entered or imported using the ADI Laboratory Report Default template.

## *4.7 STARHS Assay (Required for eHARS)*

| **Variable name and format:** | **LOINC_CD (ASSAY)** |
|---|---|
| Wording in IVR Database: | Assay |
| Wording in eHARS Lab Document: | Test |
| Values available in eHARS | ST-001 = BED<br>ST-002 = Vironostika<br>ST-003 = Avidity<br>ST-888 = Other<br>ST-999 = Unknown (Blank in IVR database) |

### *Description*

For eHARS, this variable is used to designate what lab test is being reported. For incidence surveillance, this variable, *STARHS Assay*, designates the assay used for the STARHS algorithm. The assay type is designated by a LOINC code. This variable is **required** to enter any STARHS specimen information in eHARS.

### *Purpose*

In eHARS, STARHS specimen and results information are entered like any other laboratory test; therefore, eHARS requires the selection of one of the codes for STARHS tests when entering HIS data. For HIS, the purpose of this variable is to distinguish various STARHS testing algorithms that have been used. Until March 2005, HIS used the Less Sensitive (LS) Enzyme Immunoassay (EIA) test for STARHS. Currently, HIS uses the BED HIV-1 Capture Enzyme Immunoassay (BED) for STARHS. In the future, other tests, such as an avidity test, might be used.

### *Sources*

This variable was imported with the STARHS laboratory results into the IVR database. Because the only current STARHS assay is the BED, the value of ST-001 can imported or manually entered from a drop-down menu as STARHS (BED).

### *History*

The value of ST-999 (STARHS Unknown) was used during the conversion of data into eHARS for any STARHS result that had a blank value. **Note: Do not use STARHS (Unknown), when entering information for a specimen that has not yet been sent for STARHS testing or does not yet have a STARHS result.** Use STARHS (BED). Most SAS programs exclude non-BED results, so the entered information entered as STARHS (Unknown) will be lost in data analysis.

### *Data Entry*

These data may be manually entered or imported using the ADI Laboratory Report Default template.

### *4.8 Specimen Type (Recommended for HIS)*

| **Variable name and format:** | **SPECIMEN (LSPECTY)** |
|---|---|
| Wording in IVR Database: | Specimen Type |
| | Values:<br>'1-blood finger stick'<br>'2-blood venipuncture'<br>'3-blood spot'<br>'4-oral mucosal transudate'<br>'5-urine'<br>'8-other'<br>'9-unknown' |
| Wording in eHARS Lab Document: | Sample Type |
| | Values: 'BLD', 'SAL', 'URN', 'OTH', 'UNK' |

#### *Description*

The *Specimen Type* variable represents the type of specimen that was obtained from the patient, such as blood or oral fluid. In the IVR database, there were seven values that could be selected but eHARS only offers five.

#### *Purpose*

This data element is useful to ascertain if the specimen was blood (serum/plasma) and therefore eligible for BED testing. This data element is available for all HIV tests in eHARS. CDC recommends that incidence surveillance areas start using this variable in eHARS for all diagnostic or other specimens for HIV tests within 3 months of diagnosis to indicate whether or not the specimen could be used for STARHS. This information will be very useful in tracking eligible specimens and, for evaluation, to quantify the denominator for the proportion of cases with a STARHS result among those with blood or serum specimen. This is the incidence evaluation standard as stated in *Technical Guidance for HIV/AIDS Surveillance Programs*:

> At least 85% of new diagnoses of HIV infection reported for a calendar year diagnosed using a serum/plasma specimen or having a follow-up HIV-related test conducted on a serum/plasma specimen within 3 months of the diagnosis should have a specimen transported to the CDC STARHS laboratory, assessed at 12 months after the end of the diagnosis year.

#### *Sources*

Sources for this data element are paper or electronic laboratory reports or databases.

#### *Data Entry*

These data may be manually entered or imported using the ADI Laboratory Report Default template.

### 4.9 STARHS ID (Required for eHARS and HIS)

| **Variable name and format:** | **STARHS_SAMPLE_ID (SSTARHSID)** |
|---|---|
| Wording in IVR Database: | StarhsID |
| Wording in eHARS Lab Document:: | STARHS ID (Specimen) |

#### *Description*

The *STARHS ID* variable is the identifier assigned to a specimen in order to be STARHS tested. There is only one STARHS ID assigned to a specimen that is sent for testing. If the specimen cannot be tested (e.g., due to insufficient quantity), the same STARHS ID cannot be re-used. An individual may have more than one specimen, each with a unique STARHS ID, even if the aliquots were from the same HIV diagnostic specimen.

#### *Purpose*

This variable is used to link specimen information with STARHS results received from the STARHS laboratory.

#### *Sources*

Sources for this variable are health departments, public health laboratories, tracking databases, and the Results file received from the STARHS laboratory.

#### *Data Entry*

These data may be manually entered or imported into eHARS using the ADI Laboratory Report Default template. Data entry staff will have to manually enter the *STARHS ID* on an eHARS laboratory document in order to link the STARHS results file, which has no STATENO, to cases in eHARS during a separate import process. Otherwise, the value can be imported using a SAS program that will combine specimen information from a tracking database, STATENO, and the STARHS Results file received from the STARHS laboratory.

### 4.10 Standard Optical Density

| **Variable name and format:** | **RESULT (SOD)** |
|---|---|
| Wording in IVR Database: | SOD |
| Wording in eHARS Lab Document: | Standard Optical Density |

#### *Description*

The *Standard Optical Density* variable represents the specific value of SOD found by STARHS testing.

#### *Purpose*

This variable is used to substantiate the final result interpretation described in the data element below, *Final STARHS Result*.

### *Sources*

The only source for this data element is the STARHS result file received from the STARHS laboratory.

### *Data Entry*

CDC strongly recommends that this variable be imported into eHARS using the ADI Laboratory Report Default template to avoid errors.

## *4.11 Final STARHS Result (Required for HIS)*

| **Variable name and format:** | **RESULT_INTERPRETATION (RESULT)** |
|---|---|
| Wording in IVR Database: | Lab Result |
| Wording in eHARS Lab Document: | Final Result |
| | Values:<br>01 - Long term<br>02 – Recent<br>91 – Quantity not sufficient<br>92 – Specimen never received<br>93 – Broken in transit<br>94 – Other, indeterminate<br>95 – Not sufficient antibodies<br>99 – Undefined result |

### *Description*

The *Final STARHS Result* variable represents the main BED testing result. It also contains values to explain why a STARHS specimen does have a STARHS result.

### *Purpose*

This data element is used to classify newly-diagnosed HIV infections as recent or long-standing (long term). This is one of the most important variables for HIS and is required for specimens that receive STARHS testing.

If the result of STARHS testing is not recent or long term, a reason for not testing is reported by the STARHS laboratory and represented by the codes 91–99 described above. These codes are used exclusively by the STARHS laboratory. Reasons for not shipping specimens to the STARHS laboratory are captured in *Reason Specimen Not Sent for STARHS*, which is described later in this document.

### *Sources*

The only source for this data element is the STARHS result file received from the STARHS laboratory.

### *Data Entry*

CDC strongly recommends that this variable be imported into eHARS using the ADI Laboratory Report Default template to avoid errors. Do NOT manually enter a reason for not testing in the eHARS *Final STARHS Result* field.

### *4.12 Reason Specimen Not Sent for STARHS*

| **Variable name and format:** | **SREASON** |
|---|---|
| Wording in IVR Database: | Reason Specimen Not Tested |
| Wording in eHARS Lab Document: | Reason Specimen Not Sent for STARHS |

#### *Description*

The *Reason Specimen Not Sent for STARHS* variable represents the reason why a specimen was not sent to the STARHS laboratory for STARHS testing. In the IVR database, this variable also represented the reason why the STARHS laboratory did not have a result, and was imported from the STARHS result file received from the STARHS laboratory. In eHARS, those two types of reasons will be represented by separate variables, *Reason Specimen Not Sent for STARHS* and *Final STARHS Result*.

#### *Purpose*

This data element is to explain and track reasons why the specimen was not sent for STARHS testing. This is useful for resolving STARHS specimens without results, monitoring problems with laboratories that send specimens, and evaluation of the incidence surveillance area's effort in obtaining specimens for STARHS testing.

#### *Data Entry*

Information on why a specific specimen was not sent to the STARHS laboratory for STARHS testing should be manually entered in eHARS. Do not assume, without information, that a specimen had insufficient quantity and that is why a laboratory did not send the specimen to the public health laboratory or STARHS laboratory. Only enter a reason when there is sufficient information to support it.

## 5. IVR Lab Variables Not Included in eHARS

### *5.1 Specimen Approved for STARHS*

| **Variable name and format:** | **(LAPPRVE)[2] Yes/No/Don't Know** |
|---|---|
| Wording in IVR Database: | Approved for STARHS |
| Wording in eHARS Lab Document: | Not in eHARS |

[2] IVR database variable

#### *Description*

The *Approved for STARHS* variable designated that a specific specimen was eligible for STARHS testing, either because it was collected within 3 months of date of HIV diagnosis or consent was obtained. In the macro that assembled the monthly data transfer, CDC used the variable to select which STARHS specimen result would be transferred to CDC until July 2010.

#### *Purpose*

This variable is no longer used by CDC and it is not in eHARS.

### 5.2 State Lab ID

| Variable name and format: | (SSTATELID)[2] CLIA code for State Laboratory |
|---|---|
| Wording in IVR Database: | State Lab ID |
| Wording in eHARS Lab Document: | Not in eHARS |

[2] IVR database variable

#### Description

The *State Lab ID* variable was used locally to link specimens to a public health laboratory.

#### Purpose

This variable was never used by CDC and it is not in eHARS.

### 5.3 HIV Diagnosis Test Type

| Variable name and format: | (LTSTYP)[2] |
|---|---|
| Wording in IVR Database: | State Lab ID |
| | Values:<br>5220-9 = EIA / Elisa<br>21009-6 = Western Blot<br>5472-6 = CD4<br>25835-0 = Viral Load (NASBA)<br>5017-9 = Viral Load (bDNA)<br>25836-8 = Viral Load (RT-PCR) |
| Wording in eHARS Lab Document: | Not in eHARS |

[2] IVR database variable

#### Description

The *HIV Diagnosis Test Type* variable was used in the IVR database to indicate the type of confirmatory HIV test from which the STARHS specimen might be obtained.

#### Purpose

This variable, used locally for tracking potential STARHS specimens, was never used by CDC and it is not in eHARS. All HIV tests are entered separately into eHARS ACRF or Lab documents and can be found in the Document datasets.

### 5.4 Results Received

| Variable name and format: | (RESULTRECEIVED)[2] Yes/No |
|---|---|
| Wording in IVR Database: | Result Received |
| Wording in eHARS Lab Document: | Not in eHARS |

[2] IVR database variable

### *Description*

The *Results Received* variable was used in the IVR database to indicate that a specimen had a STARHS result imported into the database.

### *Purpose*

This variable, used for monitoring STARHS result, was never used by CDC and it is not in eHARS.

## Appendices

*Appendix A: Quick Reference for Data Analysts*

*Appendix B: List of ARV Medications for TTH*

*Appendix C: Data Entry Recommendations for HIV Incidence Surveillance Data Elements in eHARS*

Appendix A

## Quick Reference for Data Analysts

### Table A.1: Testing and Treatment History Document in HIS and eHARS Datasets

| IVR Screen Question Number[1] | | eHARS Screen Label on TTH Document | HIS Dataset Before Conversion | | | eHARS Document Dataset After Conversion | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Variable Name (Pre/Post) | Valid Values | SAS Format | Variable Name | Valid Values | Length | SAS Format[2] | SAS Label[3] |
| Pre | Post | | | | | | | | | |
| 1 | 1 | 1. Main source of testing and treatment history information | UCTS/ KCTS | 0- No<br>1-Yes<br>2-Patient Interview<br>3-Medical Record Review<br>4-From PEMS<br>5-Other | $50 | UCTS | 1-Provider Report<br>2-Patient Interview<br>3-Medical Record Review<br>4-NHM&E/PEMS<br>5-Other | $1 | $2 | Main source of TTH |
| 1 | 1 | 2. Date patient reported information | UQINTD/ KQINTD | Date | yyyymmdd<br>$8 | UQINTD | mm/dd/yyyy | $10 | yyyymmdd | Date patient reported info |
| 4a | N/A | 3. Ever had a previous positive HIV test? | UPASTP | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | UPASTP | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | Ever previous positive test |
| 4b | 3 | 4. Date of first positive HIV test | UFPOSD/ KFPOSD | Date | yyyymmdd<br>$8 | UFPOSD | mm/dd/yyyy | $10 | yyyymmdd | Date first positive test |
| 4f | 8 | 5. Ever had a negative HIV test? | UNGTST/ KNGTST | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | UNGTST | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | Ever tested negative |
| 4f | 8a | 6. Date of last negative HIV test | ULSTND/ KLSTND | Date | yyyymmdd<br>$8 | ULSTND | mm/dd/yyyy | $10 | yyyymmdd | Date last negative test |
| 5b | 9 | 7. Number of negative HIV tests within 24 months before first positive test (Do not include first positive HIV test) | UNUMTSTS/ KNUMTSTS | 1 to 99<br>R=Refused<br>D=Don't know | $2 | UNUMTSTS | 0 to 99<br>R=Refused<br>D=Don't know | $2 | $2, restricted to integer 1-99 or R or D | Number negative tests 24 mos. before first positive |
| 7 | 10 | 8. Ever taken any antiretroviral medications (ARVs)? | UHRT/ KHRT | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | UHRT | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | Ever taken any ARVs |
| 7a | 10a | 9. If yes, name(s) of ARV medication taken | PREMED1-N/ POSTMED1-N[4] | Same as Appendix B, not including 32-36 | $2 | UHRTA1 | See Appendix B | $256 | $DRUG (piped values) | Names of ARVs taken |
| 7b | 10b | 10. Dates ARVs taken: Date first began | UHRTBD/ KHRTBD | Date | yyyymmdd<br>$8 | UHRTBD | mm/dd/yyyy | $10 | yyyymmdd | Date first began ARVs |

| IVR Screen Question Number[1] | | eHARS Screen Label on TTH Document | HIS Dataset Before Conversion | | | eHARS Document Dataset After Conversion | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Variable Name (Pre/Post) | Valid Values | SAS Format | Variable Name | Valid Values | Length | SAS Format[2] | SAS Label[3] |
| Pre | Post | | | | | | | | | |
| 7d | 10d | 11. Date of last ARV use | UHRTED/ KHRTED | Date | yyyymmdd $8 | UHRTED | mm/dd/yyyy | $10 | yyyymmdd | Date last ARV use |
| **Optional/Legacy Variables** | | | | | | | | | | |
| 7c | 10c | 12. Are you now taking any ARVs? | QHRTNW/ KHRTNW | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | QHRTNW | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | Now taking ARVs |
| 4 | N/A | 13. Ever been tested for HIV before today? (Legacy Pre-test only) | UPTESTS | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | UPTESTS | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | Ever tested for HIV before |
| 6 | 7 | 14. When was the first time you ever got tested for HIV? | UFTSTD/ KFTSTD | Date | yyyymmdd $8 | UFTSTD | mm/dd/yyyy | $10 | yyyymmdd | Date first HIV test |
| 4c | 4 | 15. When you first tested positive for HIV, was the HIV test an anonymous test? | UFPOSA/ KFPOSA | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | UFPOSA | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | Was first positive test anonymous |
| 5a | N/A | 16. In the two years before first positive test, how many times did you get tested for HIV? | UPNUMTSTS | 1 to 99<br>R=Refused<br>D=Don't know | $2 | UPNUMTSTS | 0 to 99<br>R=Refused<br>D= Don't know | $2 | $2, restricted to integer 1-99 or R or D | Number of tests 2 yrs before previous positive (legacy pre-test) |
| 3a | N/A | 17. Reasons for getting today's HIV test (Legacy Pre-test form only): a. Think you might have been exposed to HIV in the 6 months before the test | UREAS3_1 | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | UREAS3_1 | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | Today's test reason: exposed past 6 mos. |
| 3b | N/A | b. Get tested on a regular basis and it is time to get tested again | UREAS3_2 | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | UREAS3_2 | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | Today's test reason: tests regularly |
| 3c | N/A | c. Just checking to make sure you are HIV negative | UREAS3_3 | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1. | UREAS3_3 | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | Today's test reason: checking if negative |
| 3d | N/A | d. Required to get the test by insurance, military, court, or other agency | UREAS3_4 | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1. | UREAS3_4 | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | Today's test reason: required by agency |

| IVR Screen Question Number[1] | | eHARS Screen Label on TTH Document | HIS Dataset Before Conversion | | | eHARS Document Dataset After Conversion | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Variable Name (Pre/Post) | Valid Values | SAS Format | Variable Name | Valid Values | Length | SAS Format[2] | SAS Label[3] |
| Pre | Post | | | | | | | | | |
| 3e | N/A | e. Other reason you wanted to get tested | UREAS3_5 | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | UREAS3_5 | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | Today's test reason: other |
| 3e | N/A | f. If other reason, describe: | UR3_5SP | Text field | $50 | UR3_5SP | Text | $50 | text field | Today's test reason: describe |
| 4e(a) | 6a | 18. Reason for getting the first positive HIV test: a. Thought you might have been exposed to HIV in the 6 months before the test | URS4E_1/ KREAS6_1 | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | URS4E_1 | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | First positive reason: exposed past 6 mos. |
| 4e(b) | 6b | b. Got tested on a regular basis and it was time to get tested again | URS4E_2/ KREAS6_2 | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | URS4E_2 | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | First positive reason: tests regularly |
| 4e(c) | 6c | c. Just checking to make sure you were HIV negative | URS4E_3/ KREAS6_3 | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | URS4E_3 | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | First positive reason: checking if negative |
| 4e(d) | 6d | d. Required to get the test by insurance, military, court, or other agency | URS4E_4/ KREAS6_4 | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | URS4E_4 | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | First positive reason: required by agency |
| 4e(e) | 6e | e. Other reason you wanted to get tested | URS4E_5/ KREAS6_5 | 1=Yes<br>0=No<br>7=Refused<br>9=Don't know | $1 | URS4E_5 | Y=Yes<br>N=No<br>R=Refused<br>D=Don't know | $1 | $YNRD | First positive reason: other |
| 4e(e) | 6e | f. If other reason, describe: | URS4E_5SP/ KReas6_5SP | Text field | $50 | URS4E_5SP | Text | $50 | $50 | First positive reason: describe |
| 4d | 5 | 19. Name of facility where first tested positive for HIV | UFPSTYP_SITE/ KFPS_SITE | Text field | $50 | UFPS_SITE | Text | $50 | $50. | Facility name where first tested positive |
| 4d | 5 | 20. State of facility where first tested positive for HIV | UFPSTYP_STATE/ KFPS_STATE | 2-digit State codes | $2 | UFPS_STATE | 2-digit State codes | $2 | $STATE | State where first tested positive |
| 4d | 5 | 21. Type of facility where first tested positive for HIV | UFPSTYP/ KFPS | Same as eHARS Facility codes[2] | $7 | UFPSTYP | eHARS Facility codes[2] | $7 | $FAC_TYP | Facility type where first tested positive |

| IVR Screen Question Number[1] | | eHARS Screen Label on TTH Document | HIS Dataset Before Conversion | | | eHARS Document Dataset After Conversion | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Pre | Post | | Variable Name (Pre/Post) | Valid Values | SAS Format | Variable Name | Valid Values | Length | SAS Format[2] | SAS Label[3] |
| 4g | 8b | 22. Name of facility where last tested negative for HIV | ULSTNGS_SITE/ KLSTNGS_SITE | Text field | $50 | ULSTNGS_SITE | Text | $50 | $50. | Facility name where last tested negative |
| 4g | 8b | 23. State of facility where last tested negative for HIV | ULSTNGS_STATE/ KLSTNGS_STATE | 2-digit State codes | $2 | ULSTNGS_STATE | 2-digit State codes | $2 | $STATE | State where last tested negative |
| 4g | 8b | 24. Type of facility where last tested negative for HIV | ULSTNGS/ KLSTNGS | Same as eHARS Facility codes[2] | $7 | ULSTNGS | eHARS Facility codes[2] | $7 | $FAC_TYP | Facility type where last tested negative |

[1]Where a question has no number in the IVR database, the approximate position is given

[2]For SAS formats, see eHARS Technical Reference Guide, Chapter 17

[3]SAS labels not included in eHARS version 3.2

[4]The dataset may contain up to N medications for a given TTH; each will be stored in a separate variable

## Table A.2: Consent Document in HIS and eHARS Datasets

| eHARS Screen Label on Consent Document | HIS Dataset Before Conversion | | | eHARS Document Dataset After Conversion | | | | |
|---|---|---|---|---|---|---|---|---|
| | Variable Name | Valid Values | SAS Format | Variable Name | Valid Values | Length | SAS Format[1] | SAS Label |
| Date of first approach for Consent | C1DATE | Date | yyyymmdd $10 | CDATE1 | mm/dd/yyyy | $8 | $8 | cdate1 |
| Did the participant consent on this date? | CCONSENT1 | 1–Yes<br>2–No | $1 | CCONSENT1 | Y-Yes<br>N-No<br>U-Unknown | $1 | $YNU | cconsent1 |
| What visit was this? | CCONSENTVISIT | 1–Pre-test<br>2–Post-test<br>3–Other follow-up | 11. | CCONSENTVISIT1 | 01–Pre-test<br>02–Post-test<br>03–Other follow-up | $2 | $CONSENT | cconsentvisit1 |
| Date of second approach, if applicable | C2DATE | Date | yyyymmdd $10 | CDATE2 | mm/dd/yyyy | $8 | $8 | cdate2 |
| Did the participant consent on this date? | CCONSENT2 | 1–Yes<br>2–No | $1 | CCONSENT2 | Y-Yes<br>N-No<br>U-Unknown | $1 | $YNU | cconsent2 |
| What visit was this? | CCONSENTVISIT2 | 2–Post-test<br>3–Other follow-up | 11. | CCONSENTVISIT2 | 01–Pre-test<br>02–Post-test<br>03–Other follow-up | $2 | $CONSENT | cconsentvisit2 |

[1]For SAS formats, see eHARS Technical Reference Guide, Chapter 17

## Table A.3: STARHS Specimen and Result Information in HIS and eHARS Datasets

| | | HIS Dataset Before Conversion | | | eHARS Document Dataset After Conversion | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **IVR Screen Label (in order of appearance in eHARS)** | **eHARS Screen Label on Lab Document** | **Variable Name[1]** | **Valid Values** | **SAS Format** | **Variable Name** | **Valid Values** | **Length** | **SAS Format[2]** | **SAS Label** |
| Test Lab ID | Lab Name | LABID1−3 | 21D0649758<br>33D0654341<br>50D0661430 | CLIA Code[3]<br>$11 | CLIA_UID | CLIA laboratory codes[3] on CLIA_CODE table<br>Current: 33D2005937 | $11 | $11 | CLIA_UID |
| Source Lab Specimen ID | Sample ID | LSRCEID1−3 | Text | $20 | SAMPLE_ID | Text | $50 | $50 | SAMPLE_ID |
| State Lab Specimen ID | Accession Number | SSTATEID1−3 | Text | $20 | ACCESSION_NUMBER | Text | $50 | $50 | ACCESSION_NO |
| Date of Specimen Collection | Collection Date | LDTEOBT1−3 | Date | yyyymmdd<br>$8 | SAMPLE_DT | mm/dd/yyyy | $8 | yyyymmdd<br>$8 | SAMPLE_DT |
| Test Date | Result Date | TESTDATE1−3 | Date | yyyymmdd<br>$8 | RESULT_RPT_DT | mm/dd/yyyy | $8 | yyyymmdd<br>$8 | RESULT_RPT_DT |
| Not in IVR | Received Date | | | | RECEIVE_DT | mm/dd/yyyy | $8 | yyyymmdd<br>$8 | RECEIVE_DT |
| Assay | Test | ASSAY1−3 | BED=BED<br>BVLS=BVLS (Vironostika LS)<br>OTLS=OTLS (Vironostika LS)<br>OTV=OTV (Vironostika LS)<br>AVID=AVID | $7 | LOINC_CD | LOINC codes for lab tests on LOINC_CODE table | $7 | $LNC_CD | LOINC_CD |
| Specimen Type | Sample Type | LSPECTY1−3 | 1-Blood Finger Stick<br>2-Blood Venipuncture<br>3-Blood Spot<br>4-Oral Mucosal Transudate<br>5-Urine<br>8-Other<br>9-Unknown | $15 | SPECIMEN | BLD−Blood<br>OTH−Other<br>SAL− Saliva<br>URN−Urine<br>UNK−Unknown | $3 | $SPC_TYP | SPECIMEN |
| SStarhsID | STARHS ID (Specimen) | SSTARHSID1−3 | | $15 | STARHS_SAMPLE_ID | Text | $15 | $15 | STARHS_SAMPLE_ID |
| Test SOD | Standard Optical Density | SOD1−3 | | $10 | RESULT | #.### | $10 | $10 | RESULT |

| IVR Screen Label (in order of appearance in eHARS) | eHARS Screen Label on Lab Document | HIS Dataset Before Conversion | | | eHARS Document Dataset After Conversion | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Variable Name[1] | Valid Values | SAS Format | Variable Name | Valid Values | Length | SAS Format[2] | SAS Label |
| Test Result | Final Result | RESULT1–3 | 01–Long term<br>02–Recent<br>91–QNS<br>92–Not rec'd by STARHS Lab<br>93–Broken<br>94–Other | $2 | RESULT_INTERPRETATION | 01–Long term<br>02–Recent<br>91–Quantity not sufficient<br>92–Specimen never received<br>93–Broken in transit<br>94–Other, indeterminate<br>95–Not sufficient antibodies<br>99–Undefined result | $100 | $ST_RSLT | RESULT_INTERPRETATION |
| Reason Specimen Not Tested | Reason Specimen Not Sent for STARHS | SREASON1–3 | 1–QNS<br>2–Specimen never rec'd at public lab<br>3–Broken in transit<br>4–Other | $1 | SREASON | 1–Quantity not sufficient<br>2–Specimen never received at public lab<br>3–Specimen broken in transit<br>4–Other<br>5–Not sufficient antibodies | $1 | $ST_RSN | SREASON |
| Source Lab CLIA code | Not in eHARS | LSRCEL1–3 | Laboratory CLIA code[3] | $20 | Not in eHARS | | | | |
| State Lab ID | Not in eHARS | SSTATELID1–3 | State Laboratory CLIA code[3] | $20 | Not in eHARS | | | | |
| HIV Test | Not in eHARS | LTSTTYP | 5220-9=EIA/Elisa<br>21009-6=Western Blot<br>5472-6=CD4<br>25835-0=Viral Load (NASBA)<br>5017-9=Viral Load (bDNA)<br>25836-8=Viral Load (RT-PCR) | $15 | Not in eHARS | | | | |
| Approved for STARHS | Not in eHARS | LAPPRVE1–3 | 0=no<br>1=yes<br>2=pending | $1 | Not in eHARS | | | | |

[1]HIS datasets include specimen information and STARHS results for up to 3 specimens; variables are numbered for each corresponding sample.

[2]For SAS formats, see eHARS Technical Reference Guide, Chapter 17

[3]To look up specific CLIA codes, refer to: http://www.hipaaspace.com/Medical_Billing/Coding/Clinical_Laboratory_Improvement_Amendments/CLIA_Codes_Lookup.aspx

Appendix B

## List of ARV Medications for TTH

Medications appear alphabetically in eHARS and are stored in the database using the numeric codes listed. More than one medication can be selected by holding down the Control key. The variable UHRTA1 in the document-based datasets consolidates all the medications selected in a document in a piped fashion. For example, if the person reported taking Ziagen, lamivudine, and zidovudine, the values would appear as '20|03|26'.

| Name of Medication | Code |
|---|---|
| Agenerase (amprenavir) | 22 |
| Aptivus (tipranavir, TPV) | 30 |
| Atripla (efavirenz / emtricitabine / tenofovir DF) | 32 |
| Combivir (lamivudine/ zidovudine) | 24 |
| Crixivan (indinavir sulfate) | 06 |
| Emtriva (emtricitabine, FTC) | 11 |
| Epivir (lamivudine, 3TC) | 03 |
| Epzicom (3TC / ABC) | 28 |
| Fortovase (saquinavir) | 25 |
| Fuzeon (enfuvirtide, T-20) | 10 |
| Hepsera (adefovir) | 19 |
| Hivid (zalcitabine, ddC) | 02 |
| Hydroxyurea | 23 |
| Invirase (saquinavir mesylate) | 18 |
| Intelence (etravirine) | 34 |
| Isentress (raltegravir) | 36 |
| Kaletra (lopinavir / ritonavir) | 16 |
| Lexiva (fosamprenavir, 908) | 31 |
| Norvir (ritonavir) | 07 |
| Other | 88 |
| Prezista (darunavir, DRV) | 33 |
| Rescriptor (delavirdine mesylate) | 09 |
| Retrovir (zidovudine, ZDV, AZT) | 26 |

| Name of Medication | Code |
|---|---|
| Reyataz (atazanavir sulfate) | 15 |
| Saquinavir (Fortavase, Invirase) | 08 |
| Selzentry (maraviroc) | 35 |
| Sustiva (efavirenz) | 21 |
| Trizivir (abacavir sulfate / lamivudine / zidovudine) | 13 |
| Truvada (FTC / TDF) | 27 |
| Unspecified | 99 |
| Videx (didanosine, ddl) | 01 |
| Videx EC (didanosine, ddl) | 14 |
| Viracept (nelfinavir mesylate) | 17 |
| Viramune (nevirapine) | 05 |
| Viread (tenofovir) | 12 |
| Zerit (stavudine, d4T) | 04 |
| Ziagen (abacavir sulfate) | 20 |

Appendix C

## Data Entry Recommendations for HIV Incidence Surveillance Data Elements in eHARS

### Overall Guidance

Data entry in eHARS is document-based. Document-based data entry means that data entered or imported on an individual document in eHARS were collected at the same time and from the same source. Documents in eHARS that relate to HIV Incidence Surveillance (HIS) include the Testing and Treatment History (TTH) Document and the Laboratory Document (where specimen information and STARHS results are entered).

On the TTH and Laboratory documents in eHARS, there are a number of information tabs which contain many data fields. These are the tabs (in italics) and their descriptions for TTH and Laboratory documents:

**Testing and Treatment History** document:

1. *Form Info* – includes Stateno, reporting city, source of document, surveillance method, etc.
2. *Identification* – includes name, address, identification numbers
3. *Demographics* – includes DOB, sex, marital status, education, ethnicity, race, country of birth, etc.
4. *Testing and Treatment Information* – includes all required and optional TTH data elements
5. *Local Fields* – fields that can be added by eHARS Administrator, for local use, not sent to CDC
6. *Comments* – text field for extra information for local use, not sent to CDC

**Laboratory** document:

1. *Form Info* – includes Stateno, reporting city, source of document, surveillance method, etc.
2. *Identification* – includes name, address, identification numbers
3. *Demographics* – includes DOB, sex, marital status, ethnicity, race
4. *History* – includes risk factors for HIV
5. *Lab Data* – STARHS lab name, sample ID, collection date, test type, STARHS ID, STARHS results, etc.
6. *Local Fields* – fields that can be added by eHARS Administrator, for local use, not sent to CDC
7. *Comments* – text field for extra information for local use, not sent to CDC

### Data Entry in eHARS

When entering or importing data into a document in eHARS, only data contained on a given data collection form should be entered in that document. For example, if demographic information, such as date of birth, race or ethnicity, is included on a TTH data collection form, these demographic variables should be entered on the *Demographics* tab of the TTH document. Similarly, if name appears on the data collection form, it should be entered on the *Identification* tab. However, if the TTH data collection form does not contain any identification or demographics fields or if the fields are blank, then do not enter anything on the *Identification* or *Demographics* tabs and do not seek this information from another document in eHARS. Furthermore, if these data are available from another source, they should be entered on a separate eHARS document. Note: demographic and identification information entered on eHARS TTH and Lab documents can affect Person View calculations, for example race, ethnicity, and address, depending on document hierarchies.

Many surveillance areas use an ACRF form that contains a section for HIV Testing and Antiretroviral Use History. In this case, data for the ACRF should be entered on an ACRF document and the supplemental TTH information on the ACRF should be entered on a separate TTH document in

eHARS. It is not necessary to re-enter the identification or demographic information on the *Identification* and *Demographics* tabs on the eHARS TTH document as these data will have been entered on the eHARS ACRF document.

There is one tab that is very important for all eHARS documents, the *Form Info* tab. All documents in eHARS include a minimal number of required/recommended data elements about the document which need to be completed on the *Form Info* tab, as recommended below for incidence-related documents. These are used for evaluation and data quality purposes.

**Required, Recommended and Optional Data Elements**
The following clarifies the required, recommended, and optional data elements that appear on incidence-related eHARS documents. In this document, 'Required' data elements are those that are important for incidence estimation or that are necessary to enter an incidence document in eHARS. 'Recommended' data elements are very important for evaluation or matching purposes, but may not always be available. 'Optional' variables are used for local surveillance and evaluation needs or for enhancing case information. Tabs or fields found in eHARS that are not listed below are optional and can be left blank.

## TESTING AND TREATMENT HISTORY DOCUMENT

### *Form Info tab*

| | |
|---|---|
| Stateno | (Required) |
| Reporting City | (Optional) |
| Date Received at Health Dept. | (Optional) |
| Source | (Recommended) – location where data were obtained |
| Surveillance Method | (Recommended) – used for evaluation |
| Report Medium | (Optional) |
| Date Form Completed | (Optional) – recommended if this appears on the form |
| Person Completing Form | (Optional) |
| Facility completing form | (Optional) |

### *Identification tab* *[Enter only if this information is on the data collection form]*

Name type
First Name (Recommended for matching purposes)
Middle Name (Recommended for matching purposes)Last Name
(Recommended for matching purposes)
Address information
ID

### *Demographics tab* *[Enter only if this information is on the data collection form]*

Sex at Birth
Date of Birth (Recommended for matching purposes)
Ethnicity
Race
Country of Birth

### *Testing and Treatment Information tab*

**Required Data Elements** (leave blank if no attempt was made to obtain information)**:**

1. Main source of testing and treatment history information
2. Date patient reported information
3. Ever had a previous positive HIV test?
4. Date of first positive HIV test
5. Ever had a negative HIV test?
6. Date of last negative HIV test
7. Number of negative HIV tests within 24 months before first positive test (Do not include first positive HIV test)
8. Ever taken any antiretroviral medications (ARVs)?
9. If yes, name(s) of ARV medication taken
10. Dates ARVs taken: Date first began
11. Date of last ARV use

## LAB DOCUMENT – STARHS RESULTS

### *Form Info tab*

| | |
|---|---|
| Stateno | (Required) |
| Date Received at Health Department | (Optional) |
| Source | (Recommended) – should be 'A05' |
| Did this report initiate a new investigation? | (Optional) – should be 'no' |
| Reporting Medium | (Optional) |
| Surveillance Method | (Recommended) – for evaluation, should be 'Follow-up' |

### ***Identification tab*** *[Enter only if this information is on the data collection form]*

| | |
|---|---|
| Name type | |
| First Name | (Recommended for matching purposes) |
| Middle Name | (Recommended for matching purposes) |
| Last Name | (Recommended for matching purposes) |

### *Lab Data tab*

| | |
|---|---|
| Lab Name | (Recommended) |
| Sample ID (Specimen) | (Recommended) |
| Accession Number | (Recommended) |
| Collection Date | (Required for STARHS specimen document) |
| Result Date | (Recommended for valid BED results) |
| Received Date | (N/A) |
| Test (test type) | (Required) – necessary to display fields for BED test information |
| Sample Type | (Recommended) |
| STARHS ID (Specimen) | (Required) |
| Final Result | (Required if available) |
| Standard Optical Density | (Required with valid STARHS Final Result) |
| Reason Specimen Not Sent for STARHS | (Optional) |

# Local HIV Incidence Estimation Guide

| | |
|---|---|
| **Version:** | 2.2 |
| **Date:** | January 24, 2011 |
| **Prepared by:** | Rebecca Ziebell<br>Qian An<br>Xing Dong |

# Version Tracking

| Version | Date Finalized | Comments |
|---|---|---|
| 2.0 | October 21, 2010 | Initial version of user guide for 2nd round of local HIV incidence estimation |
| 2.1 | January 7, 2011 | Updates to subsection 2.2 and sections 3 and 4 |
| 2.2 | January 24, 2011 | Updates to data dictionary table in subsection 3.1 |

# Table of Contents

1. Background ..... 4
2. Required Materials ..... 4
2.1. Locally Available Materials ..... 4
2.2 CDC-Provided Materials ..... 4
3. HIV Incidence Estimation ..... 5
3.1. PGM1_PREPARE_DATA ..... 6
Table 1. Assess Missingness of STARHS Results, Testing Group, and T (Among Observed Repeat Testers) ..... 12
Table 2. Assess Distribution of Observed STARHS Result Data ..... 12
Table 3. Assess Distribution of Observed Testing Group Data ..... 12
3.2. PGM2_IMPUTE_BED_GROUP_T ..... 13
Table 4. Incidence Data Summary Before and After Imputation ..... 14
Table 5. Determine Which Potential Strata Meet 20%/40/10 Requirement ..... 15
3.3. PGM3_DETERMINE_STRATA ..... 17
Table 6. Determine Stratification Variable Based on BED-Recent Heterogeneity ..... 19
3.4. PGM4_ESTIMATE_INCIDENCE ..... 20
4. Interpreting HIV Incidence Estimates ..... 24
4.1. PGM5_COMPARE_ESTIMATES ..... 24
4.2. PGM6_EVALUATE DIFFERENCES ..... 27
5. Contact Information ..... 28

# 1. Background

The purpose of this document is to serve as a reference for HIV incidence surveillance staff during preparation of local HIV incidence estimates. For more general information on the purpose of HIV incidence estimation and the history of HIV incidence surveillance in the United States, please refer to the resources available on the CDC website: http://www.cdc.gov/hiv/topics/surveillance/incidence.htm

The following acronyms and abbreviations will be used throughout this document:

| Abbreviation | Definition |
|---|---|
| ARV | Antiretroviral drug |
| BED | BED HIV-1 EIA Capture Assay |
| eHARS | Enhanced HIV/AIDS Reporting System |
| HICSB | CDC's HIV Incidence and Case Surveillance Branch |
| HIS | HIV Incidence Surveillance |
| STARHS | Serologic Testing Algorithm for Recent HIV Seroconversion |
| TTH | Testing and Treatment History |
| XX | Placeholder for abbreviation of name of local surveillance area |

# 2. Required Materials

This section details the materials needed to successfully estimate HIV incidence at the state or local level. Unless otherwise noted, the reader can assume implied file extensions of *.sas for SAS programs, *.sas7bdat for data sets, and *.pdf for output tables. Throughout this document, the names of SAS programs will appear in ALL CAPS, data set names in SMALL CAPS, variable names in *italics*, and PDF filenames in **bold**.

## 2.1. Locally Available Materials

The local HIV incidence estimation process requires that HIS staff have access to several SAS data sets exported from the local eHARS system. The latest available frozen, year-end PERSON, DOCUMENT and LAB data sets are specifically required as inputs to the local HIV incidence estimation process. The process also uses the PERSON_MI imputed risk factor data set[1], which should be based on the same frozen, year-end snapshot of the eHARS system. The final required local data set is the monthly HIS data set (e.g., XX_HIS_YYYYMMDD[2]), which need not be based on the same frozen, year-end eHARS data sets; instead, the most recent monthly HIS data set can be used for local HIV incidence estimation.

## 2.2 CDC-Provided Materials

CDC has developed a set of six SAS programs to facilitate the estimation of HIV incidence at the local level. These programs should be opened and executed in SAS 9.1.3 or SAS 9.2., not in SAS Enterprise Guide. The programs were designed to be run with a minimum of user interaction. Sections that do require user input appear immediately beneath program headers and are clearly marked as follows:

[1] PERSON_MI refers to the output of the MI SAS PROGRAMS 2009 V1 program, which was provided to core surveillance staff by CDC's Quantitative Sciences and Data Management Branch on February 4, 2010.
[2] XX refers to the two- or three-character abbreviation of the project area name; see section 3.1 for more details. YYYYMMDD represents the date of data set creation. The monthly HIS data set must be in the revised PERSON-based format implemented in late summer 2010.

```
*~*~*~
```

- Outputs HIV incidence estimation results tables

### 3.1. PGM1_PREPARE_DATA

PGM1_PREPARE_DATA begins with an interactive section in which the user must enter the desired year or range of years for which to estimate HIV incidence, the local area abbreviation, and the names and locations of the input data sets described in sections 2.1 and 2.2 above. This introductory section is clearly noted in the program and uses the "%let" style of setting macro variables (which should be familiar to individuals who have used other SAS programs provided by HICSB's Incidence and Viral Resistance Team):

```
*~*~*~
```

- Multi-word name: First letter of each word (Los Angeles County = LAC, New York City = NYC, San Francisco = SF)
- Single-word name: First three letters of name (Chicago = CHI, Houston = HOU, Philadelphia = PHI)

After entering the macro variable values, the user can click the Submit button to execute PGM1_PREPARE_DATA. The program includes a check to ensure that the baseline data set contains sufficient data to estimate HIV incidence for the desired year(s). If local data do not meet the minimum requirements—i.e., for each year, 300 diagnoses of HIV disease and 15% STARHS result completeness—then the program will stop executing and display an alert window. The window alerts the user to the specific year or years that lack sufficient data. At this point the user can return to the interactive portion at the beginning of the program and enter a different year or years for which to estimate HIV incidence, then resubmit PGM1_PREPARE_DATA.

Successful execution of PGM1_PREPARE_DATA accomplishes the following:

- Identification of the baseline sample population—i.e., adult and adolescent diagnoses of HIV disease in the desired year(s), residing in the selected HIS area[3] at time of diagnosis—from the frozen, year-end eHARS PERSON data set
- Collection of available documented dates of last negative HIV tests from frozen, year-end eHARS DOCUMENT and LAB data sets
- Addition of HIS-specific data elements from latest monthly HIS data set, XX_HIS_YYYYMMDD
- Selection of "best" TTH data elements for each case:
    - Earliest date of first positive HIV test
    - Latest date of last negative HIV test
    - Earliest ARV start date
    - Latest ARV end date
    - Most complete set of Ever Tested Negative, Number of Negative HIV Tests in 24 Months Before First Positive, and Facility of First Positive Test data elements
- Elimination of non-useable STARHS results (i.e., from non-BED assays, specimens collected outside the three-month post-diagnosis timeframe, or persons who used ARV during six months prior to specimen collection)
- Resolution of apparent conflicts between TTH data elements
- Creation of various calculated variables for use in imputation and estimation programs
- Creation of the initial HIV incidence estimation data set, XX_HIS_EST_DATA_YEAR(S)[4]
- Generation of pre-imputation data review tables

[3] PGM1_PREPARE_DATA is designed to prepare an initial HIV incidence estimation data set at the level of the funded project area. In some cases—e.g., when the funded area is "State, excluding Major City"—the desired area for estimation may be something other than the area that was funded. In these cases, two or more HIS areas (e.g., State and Major City) will need to collaborate to create a statewide, deduplicated version of the monthly HIS data set to use as input for PGM1_PREPARE_DATA. Local HIS staff should contact their assigned HICSB epidemiologist for assistance.

[4] Single-year filename: XX_HIS_EST_DATA_YEAR; multi-year filename: XX_HIS_EST_DATA_STARTYEAR_ENDYEAR

The table below contains a list of pertinent calculated XX_HIS_EST_DATA_YEAR(S) variables in the order of their creation in PGM1_PREPARE_DATA. All other variables in XX_HIS_EST_DATA_YEAR(S) are sourced directly from the eHARS PERSON data set or the monthly HIS data set.

| Variable Name | Variable Label | Calculation Method & Valid Values |
|---|---|---|
| *hars_fpos_yrmo* | Date of HIV disease diagnosis | Numeric YYYYMM version of variable *hiv_aids_dx_dt* from eHARS PERSON data set |
| *hars_fpos_yr* | Year of HIV disease diagnosis | Numeric YYYY version of variable *hiv_aids_dx_dt* from eHARS PERSON data set |
| *hiv2aid* | Months from HIV diagnosis to AIDS diagnosis | Calculated from eHARS PERSON data set variables *hiv_dx_dt* and *aids_dx_dt* |
| *aids* | AIDS diagnosis at HIV diagnosis | If calculated variable *hiv2aid* is less than or equal to 0, then *aids* = "1"; otherwise, *aids* = "0". |
| *aids6* | AIDS diagnosis 1–6 months after HIV diagnosis | If calculated variable *hiv2aid* is between 1 and 6 (inclusive), then *aids6* = "1"; otherwise, *aids6* = "0". |
| *res_st* | State of residence at HIV disease diagnosis | If state of residence at HIV Dx (*rsh_state_cd*) is not missing, then state of residence at HIV disease Dx equals *rsh_state_cd*; otherwise, state of residence at HIV disease Dx equals state of residence at AIDS Dx (*rsa_state_cd*). |
| *res_county* | County of residence at HIV disease diagnosis | If state of residence at HIV Dx (*rsh_state_cd*) is not missing, then county of residence at HIV disease Dx equals county of residence at HIV Dx (*rsh_county_name*); otherwise, county of residence at HIV disease Dx equals county of residence at AIDS Dx (*rsa_county_name*). |
| *hars_geoco* | County FIPS of residence at HIV disease diagnosis | If state of residence at HIV Dx (*rsh_state_cd*) is not missing, then county FIPS of residence at HIV disease Dx equals county FIPS of residence at HIV Dx (*rsh_county_fips*); otherwise, county FIPS of residence at HIV disease Dx equals county FIPS of residence at AIDS Dx (*rsa_county_fips*). |
| *res_city* | City of residence at HIV disease diagnosis | If state of residence at HIV Dx (*rsh_state_cd*) is not missing, then city of residence at HIV disease Dx equals city of residence at HIV Dx (*rsh_city_name*); otherwise, city of residence at HIV disease Dx equals city of residence at AIDS Dx (*rsa_city_name*). |
| *hars_geocty* | City FIPS of residence at HIV disease diagnosis | If state of residence at HIV Dx (*rsh_state_cd*) is not missing, then city FIPS of residence at HIV disease Dx equals city of residence at HIV Dx (*rsh_city_fips*); otherwise, city FIPS of residence at HIV disease Dx equals city FIPS of residence at AIDS Dx (*rsa_city_fips*). |

| Variable Name | Variable Label | Calculation Method & Valid Values |
|---|---|---|
| *hars_fps* | Type of facility of HIV disease diagnosis | If type of facility of HIV Dx (*hf_facility_type_cd*) is not missing, then type of facility of HIV disease Dx equals *hf_facility_type_cd*; otherwise, type of facility of HIV disease Dx equals type of facility of AIDS Dx (*af_facility_type_cd*). Valid values: eHARS facility type codes. |
| *hars_lstn_yrmo* | Date of last negative test | Numeric YYYYMM version of latest available negative screening test result from eHARS LAB table |
| *lstnd* | TTH date of last negative test (char, yyyymmdd) | Later of pre- and post-test TTH variables *ulstnd* and *klstnd* |
| *tth_lstn_yrmo* | TTH date of last negative test (num, yyyymm) | Calculated from *lstnd* |
| *lstnd_tth_form* | TTH form (pre/post) used for date of last negative test | Valid values: “PRE”, “POST” |
| *fposd* | TTH date of first positive test (char, yyyymmdd) | Earlier of pre- and post-test TTH variables *ufposd* and *kfposd* |
| *tth_fpos_yrmo* | TTH date of first positive test (num, yyyymm) | Calculated from *fposd* |
| *fposd_tth_form* | TTH form (pre/post) used for date of first positive test | Valid values: “PRE”, “POST” |
| *hrtbd* | TTH date of start of ARV use (char, yyyymmdd) | Earlier of pre- and post-test TTH variables *uhrtbd* and *khrtbd* |
| *tth_arv_start_yrmo* | TTH date of start of ARV use (num, yyyymm) | Calculated from *hrtbd* |
| *hrtbd_tth_form* | TTH form (pre/post) used for date of start of ARV use | Valid values: “PRE”, “POST” |
| *hrted* | TTH date of end of ARV use (char, yyyymmdd) | Later of pre- and post-test TTH variables *uhrted* and *khrted* |
| *tth_arv_end_yrmo* | TTH date of end of ARV use (num, yyyymm) | Calculated from *hrted* |
| *hrted_tth_form* | TTH form (pre/post) used for date of end of ARV use | Valid values: “PRE”, “POST” |
| *other_tth_form* | TTH form (pre/post) used for key non-date variables | Pre- or post-test form chosen as source of key non-date variables based on number of useable values on each form. Valid values: “PRE”, “POST”. |
| *ngtst* | TTH ever tested negative | If calculated variable *other_tth_form* equals “PRE”, then *ngtst* equals pre-test TTH variable *ungtst*; otherwise, *ngtst* equals post-test TTH variable *kngtst*. Valid values: “0” (No), “1” (Yes), “7” (Refused), “9” (Unknown). |

| Variable Name | Variable Label | Calculation Method & Valid Values |
|---|---|---|
| *numtsts* | TTH number of tests in two years before first positive test, no previous positive (char) | If calculated variable *other_tth_form* equals "PRE", then *numtsts* equals pre-test TTH variable *unumtsts*; otherwise, *numtsts* equals post-test TTH variable *knumtsts*. |
| *pnumtsts* | TTH number of tests in two years before first/previous positive test (char) | If calculated variable *other_tth_form* equals "PRE", then *pnumtsts* equals pre-test TTH variable *upnumtsts*; otherwise, *pnumtsts* is blank. |
| *num_tests* | TTH number of tests in two years before first positive test (num) | If calculated variable *other_tth_form* equals "PRE", then *num_tests* equals the greater of pre-test TTH variables *unumtsts* and *upnumtsts*; otherwise, *num_tests* equals *knumtsts*. Note: 77 = Refused, 99 = Unknown. |
| *fps* | TTH type of facility of HIV diagnosis | If calculated variable *other_tth_form* equals "PRE", then *fps* equals pre-test TTH variable *ufpstyp*; otherwise, *fps* equals post-test TTH variable *kfps*. Valid values: eHARS facility type codes. |
| *spec_date* | Date of STARHS specimen collection (SAS date) | Calculated from specimen variable *ldteobt* |
| *raw_result* | Raw STARHS result | Unmodified/original STARHS result |
| *result* | STARHS result | Modified STARHS result, set to missing when: value of calculated variable *raw_result* is invalid (i.e., not long-term or recent); specimen was collected >3 months after HIV disease Dx date; BED assay was not used; or there is evidence of ARV use during 6 months prior to specimen collection. |
| *newdx* | Presumed new diagnosis | If calculated variable *tth_fpos_yrmo* is more than 6 months earlier than calculated variable *hars_fpos_yrmo*, then *newdx* equals "N"; otherwise, *newdx* equals "Y". |
| *cal_fpos_yrmo* | Calculated date of first positive test (num, yyyymm) | Earlier of calculated variables *hars_fpos_yrmo* and *tth_fpos_yrmo* |
| *cal_lstn_yrmo* | Calculated date of last negative test (num, yyyymm) | Later of calculated variables *hars_lstn_yrmo* and *tth_lstn_yrmo* |
| *bed* | STARHS result group | If concurrent HIV and AIDS diagnoses (*aids* = "1"), then *bed* = "a"; else if *result* is missing, then *bed* = "x"; else if STARHS result is long-term (*result* = "01"), then *bed* = "b"; else if STARHS result is recent (*result* = "02"), then *bed* = "0". |

| Variable Name | Variable Label | Calculation Method & Valid Values |
|---|---|---|
| *group* | Testing group | If calculated date of last negative test (*cal_lstn_yrmo*) is not missing and not equal to HIV disease Dx date (*hars_fpos_yrmo*) OR ever tested negative equals yes (*ngtst* = "1") OR number of tests in two years before first positive (*num_tests*) is greater than 1 and less than 77, then *group* = "1" (repeat testers); else if ever tested negative equals no (*ngtst* = "0"), then *group* = "0" (new testers); otherwise, *group* = "x" (missing). |
| *t* | Time (in months) from calculated last negative test to calculated first positive test | Difference between calculated variables *cal_lstn_yrmo* and *cal_fpos_yrmo* |
| *cal_fps* | Calculated type of facility of HIV diagnosis (raw) | If calculated date of first positive (*cal_fpos_yrmo*) equals HIV disease Dx date (*hars_fpos_yrmo*) and type of facility of HIV disease Dx (*hars_fps*) is not missing, then *cal_fps* = *hars_fps*; otherwise, *cal_fps* = TTH type of facility of HIV diagnosis (*fps*). |
| *cat_fps* | Calculated type of facility of HIV diagnosis (categorical) | Values: "1-CTS/STD", "2-Inpatient", "3-Outpatient", "4-Other", "5-Unknown" |

PGM1_PREPARE_DATA ends by generating three tables (described below), which are automatically output to the **XX Year(s) Pre-Imputation Data Review** PDF file in the "estlib" directory. These tables provide an overview of local data quality related to the imputation of missing data that occurs in the next SAS program, PGM2_IMPUTE_BED_GROUP_T. The tables are purely informational and require no direct action on the part of the user; they are simply intended to help the user understand the imputation process. Since imputation and estimation are performed separately for each year, Tables 1–3 are similarly broken out by year.

Tables 1–3 are designed to address different aspects of one of the basic assumptions of the imputation process—namely that missing data are missing-at-random (MAR). MAR means that missingness only depends on observed data points, not on the value of the unobserved data. The assumption of MAR data is impossible to prove, but most imputation models address the issue by incorporating as many potential predictors of missingness as possible so that any remaining, unexplained missing data are likely to be MAR. In accordance with this approach, the imputation model used for local HIV incidence estimation incorporates a variety of covariates that may be associated with the missingness and/or the distribution of non-missing data in the three variables being imputed—i.e., STARHS results, testing group assignments, and T. Table 1 addresses potential covariates associated with missingness of the three variables, while Tables 2 and 3 address potential covariates associated with non-missing values (which are used by the imputation model to replace missing values).

### *Table 1. Assess Missingness of STARHS Results, Testing Group, and T (Among Observed Repeat Testers)*

For each year, Table 1 displays the missingness of data to be imputed by a list of covariates: sex, race/ethnicity, age at diagnosis, raw (i.e., not imputed) transmission category, and type of facility of diagnosis. When reviewing this table, the user should pay attention to whether the percentage of cases with missing STARHS results or testing group assignments is similar across each level of a covariate (e.g., for both sexes or for all age groups). If the missing percentage differs across levels, then missingness is said to be associated with that covariate. The covariate should then be considered during imputation either by including it in the imputation model or performing a stratified imputation by the covariate.

The imputation model used in local HIV incidence estimation incorporates all of the covariates displayed in Table 1, even those that may not be associated with missingness, since including them has no negative impact on the model. Therefore, Table 1 requires no direct action on the part of the user; it is simply meant to provide insight into local data.

Table 1 should indicate low missingness of T among observed repeat testers; the imputation of T is mainly reserved for imputed repeat testers.

Some HIS areas may be aware of other covariates (e.g., county of residence) that are locally associated with missingness and should be included in the imputation model. Areas should contact their assigned HICSB epidemiologists to discuss modifying the SAS programs to include these covariates.

### *Table 2. Assess Distribution of Observed STARHS Result Data*

Table 2 is designed to highlight whether non-missing STARHS results are associated with any particular covariates. Table 2 output is broken out by year and can be reviewed separately by year as well. The covariates included in Table 2 are the same as those in Table 1: sex, race/ethnicity, age at diagnosis, raw transmission category, and type of facility of diagnosis. If the percentage of BED-recent or long-term results differs across levels of a covariate, this indicates that different types of cases have different chances of being detected as BED-recent or long-term and that the covariate in question should be included in the imputation of STARHS results. Again, the imputation model used in local HIV incidence estimation already incorporates all of the potential covariates listed in Table 2, since their inclusion has no negative impact on the model. Therefore, Table 2 requires no direct action on the part of the user; it is purely informational.

### *Table 3. Assess Distribution of Observed Testing Group Data*

Table 3 allows the user to assess whether non-missing testing group assignments and T values are associated with any covariates, which include the aforementioned demographic categories as well as observed STARHS result. Similar to Tables 1 and 2, Table 3 output is separated by year.

When reviewing Table 3, the first area of focus should be the percentage of new testers and whether it differs across levels of the potential covariates. As a reminder, new testers are those who indicated that they never had a negative test result prior to their first positive. If the percentage of new testers differs across levels of one or more covariates, this indicates that the likelihood of being a new tester differs with the covariate(s) in question, which should then be included in the imputation of missing testing

group values. Again, the local HIV incidence estimation process automatically includes all of the displayed covariates in the imputation of testing group, so no action is required on the part of the user.

The second area of focus in Table 3 is whether the median value of T among observed repeat testers is associated with any covariates. In preparing the national HIV incidence estimates, CDC analysts noticed that STARHS result and age at diagnosis are highly associated with the distribution of T. Therefore, these two covariates are included in the imputation of T in the default version of the local HIV incidence estimation programs. At the local level, Table 3 may show that other covariates are highly associated with the distribution of T; however, these associations may be the result of small sample sizes or known biases in data collection. HIS areas may choose to modify the imputation of T based on these results, but CDC does not recommend this course of action. If HIS areas elect to modify the imputation of T, they should be careful never to stratify by more than two covariates. Limiting to no more than two stratification variables ensures that each stratum contains a sufficient number of cases with observed T to prevent an overly sparse distribution from which to impute T values.

### 3.2. PGM2_IMPUTE_BED_GROUP_T

After reviewing the pre-imputation data review tables, the user should run PGM2_IMPUTE_BED_GROUP_T, which can be submitted and executed without user input. Successful execution of PGM2_IMPUTE_BED_GROUP_T accomplishes the following:

- For each year:
  - Replacement of missing transmission category values in the initial HIV incidence estimation data set, XX_HIS_EST_DATA_YEAR(S), with 20 possible values based on the imputed risk factor data set, PERSON_MI
  - Definition of dummy variables for use as covariates in multiple imputation
  - Multiple imputation of missing STARHS results and testing group assignments
  - Imputation of missing T among repeat testers based on observed distributions of T by age at diagnosis and STARHS result strata
- Creation of XX_IMPUTED_YEAR(S) data set
- Generation of post-imputation data review tables

Missing STARHS result and testing group values are imputed via multiple imputation using SAS's MI procedure. The number of imputations was set at 20 based on the percentage of missing data and the desired precision. The following basic syntax is used:

```
proc mi <options>;
  by <variables>;
  class <variables>;
  monotone (discrim <variables>);
  var <variables>;
run;
```

As previously discussed, the imputation model used in local HIV incidence estimation incorporates a variety of covariates. The control variables used in the imputation of STARHS results are race/ethnicity, age at diagnosis, imputed transmission category, type of facility of diagnosis, and observed testing

group. The covariates used in imputation of missing testing group are race/ethnicity, age at diagnosis, imputed transmission category, type of facility of diagnosis, whether AIDS was diagnosed concurrently with HIV, whether AIDS was diagnosed 1–6 months after HIV, whether STARHS result was observed or imputed, and imputed STARHS result. Furthermore, since the number of possible transmission categories differ by sex, both STARHS results and testing group assignments are imputed separately for males and females.

Missing STARHS results are imputed prior to missing testing group assignment. Imputation of missing STARHS results excludes both concurrent AIDS cases and those that were diagnosed with AIDS between one and six months after HIV, as these cases are known to represent long-term HIV infections. STARHS results from concurrent AIDS cases are not used in estimating HIV incidence, and cases that were diagnosed with AIDS between one and six months after HIV are automatically assigned BED-long-term results.

As mentioned in the discussion of Table 1 (see page 12), the user may wish to incorporate additional covariates that are known to be associated with local missingness of STARHS results and/or testing group assignments. These modifications should be made in the multiple imputation section of PGM2_IMPUTE_BED_GROUP_T. However, the variables in PROC MI's MONOTONE and VAR statements must be listed a specific order. Therefore, the local user should consult with his or her assigned HICSB epidemiologist before modifying the PROC MI syntax.

Missing values of T among observed and imputed repeat testers are imputed through multinomial distributions within each stratum determined by age at diagnosis and STARHS result. Within each stratum, the observed distribution of T among observed repeat testers is derived and used to impute missing values of T among observed and imputed repeat testers.

Program 2 ends by generating Tables 4 and 5, which are automatically output to the **XX Year(s) Post-Imputation Data Review** PDF file in the "estlib" directory. These tables should be reviewed carefully, as each may indicate the need for additional user interaction with later SAS programs.

### *Table 4. Incidence Data Summary Before and After Imputation*

The purpose of Table 4 is to confirm the successful imputation of missing STARHS result, testing group, and T values. If any errors or data anomalies occur during imputation, they will be highlighted in Table 4, which displays the following hierarchical summary statistics for each year:

- Total number of new diagnoses of HIV disease
    - Among all diagnoses, numbers and percentages of cases with and without testing group assignments
        - Among cases with testing group assignments, numbers and percentages of new and repeat testers
            - Among repeat testers, numbers and percentages of cases with and without T
    - Among all diagnoses, numbers and percentages of cases with concurrent HIV and AIDS diagnoses ("AIDS cases"), cases in which AIDS was diagnosed between one and six

months after HIV ("AIDS6 cases"), and cases that were not diagnosed with AIDS within six months of HIV

- Among non-AIDS and non-AIDS6 cases, numbers and percentages of cases with and without STARHS results
    - Among cases with STARHS results, numbers and percentages of BED-recent and long-term cases

PGM2_IMPUTE_BED_GROUP_T should impute values for all missing values of STARHS result, testing group, and T. As a result of the imputation process, certain frequencies and percentages should look different in the post-imputation data set; however, other data should remain unchanged. With that in mind, the user should look for the following characteristics in each yearly iteration of Table 4:

1. The total number of new diagnoses should remain the same before and after imputation. The imputation process should fill out missing data, not increase the number of cases in the data set.
2. After imputation, all new diagnoses should have a testing group assignment. The percentage of new diagnoses with testing group assignments in the post-imputation column should be 100%.
3. Among cases with testing group assignments, the numbers of new and repeat testers should increase due to imputation; however, the difference in percentages should not be dramatic. Small changes are expected due to the fact that missing data are imputed multiple times through independent draws from an imputation model; however, the fact that missing data are imputed from observed data should preclude any dramatic changes in the percentages of new and repeat testers.
4. Among repeat testers, all those missing T before imputation will have an imputed T after imputation. In the post-imputation column, the percentage of repeat testers with T should be 100%.
5. AIDS and AIDS6 cases are excluded from STARHS result imputation. STARHS results from concurrent HIV/AIDS cases are not used in the HIV incidence estimation model, and AIDS6 cases are assigned BED-long-term results.
6. After imputation, all non-AIDS and non-AIDS6 cases should have STARHS results. Therefore, the percentage of cases with STARHS results should be 100% in the post-imputation column.
7. Finally, among cases with STARHS results, Table 4 should show slight changes in the proportion of BED-recent results, for the same reason described in #3 above.

If the output displayed in Table 4 does not match the description above, then the user should first check the SAS log window for evidence of errors during program execution. If none are found, the user should contact his or her assigned HICSB epidemiologist for assistance in identifying underlying data anomalies or imputation errors.

### *Table 5. Determine Which Potential Strata Meet 20%/40/10 Requirement*

The HIV incidence estimation model uses stratification to account for heterogeneity caused by the fact that different groups of people have different testing behaviors. In the context of HIV incidence estimation, a stratum refers to a group of cases with presumed similar testing behavior. The model estimates incidence for each stratum individually; the total HIV incidence estimate is the sum of

estimates from all strata. To ensure stable estimates, each selected stratum should satisfy the 20%/40/10 rule or requirement[5]:

- The number of non-AIDS cases with observed STARHS results must be greater than or equal to 40 and represent at least 20% of new diagnoses; and
- The number of BED-recent results must be greater than or equal to 10.

| | **All New Diagnoses** | **With STARHS Results** | | **BED-Recent** |
|---|---|---|---|---|
| | N | N | % | N |
| **Potential stratum** | -- | ≥40 | ≥20 | ≥10 |

Table 5 allows the user to assess, by year, whether a variety of potential strata meet the 20%/40/10 rule and thereby indicates whether regrouping of strata may be necessary to generate stratified incidence estimates. The potential stratification variables displayed in Table 5 include sex, race/ethnicity, age at diagnosis, and imputed transmission category. (Note: Raw transmission category is also displayed but only to show the impact of using the PERSON_MI imputed risk factor data set; imputed transmission category should be used to assess the 20%/40/10 rule.)

Table 5 is designed to help the user prepare for any desired regrouping, which takes place at the beginning of the next SAS program, PGM3_DETERMINE_STRATA. If the displayed numbers of new diagnoses, observed STARHS results, and observed BED-recent results do not meet the 20%/40/10 requirement for a given potential stratum, the relevant numbers appear in red on the report, prompting the user to regroup, or combine, certain strata to obtain the required case counts and/or percentages. In order to compare strata-level HIV incidence estimates across years, the user must use the same regrouping methodology for all years. Therefore, Table 5 output should be reviewed across all years, and associated regrouping decisions should ensure that 20%/40/10 requirements are met for all years whenever possible.

Consider, for example, a user who is estimating local HIV incidence for the years 2006 through 2008. This user would like to generate HIV incidence estimates for injection drug users (IDU), as local prevention efforts have been particularly focused on this group. When the user examines Table 5, he notices that 2006 data are too sparse to allow IDU to meet the 20%/40/10 requirement as a separate stratum; however, the IDU stratum does meet the 20%/40/10 requirement for 2007 and 2008. This means that the user cannot generate an IDU-only HIV incidence estimate for 2006, but he could do so for 2007 and 2008. At this point, the user has two options: (1) regroup the transmission category strata to satisfy the 20%/40/10 requirement across 2006, 2007, and 2008—and lose the ability to generate an IDU-only HIV incidence estimate for the latter years; or (2) regroup the transmission category stratification variable to maintain a separate, IDU-only stratum, but note that the 2006 HIV incidence estimate for IDU should be interpreted with extreme caution, recognizing the limitations of the 2006 data when strata are configured in this way. Either option is appropriate; the choice depends on the needs of the local HIS area. (In either case, when regrouped in this way, the variable *mode3* should not

[5] Formerly known as the 200/40/10 rule or requirement.

be selected as the stratification variable for the 2006 HIV incidence estimate. Selection of the best stratification variable will be discussed further in the description of Table 6 and in section 3.4 below.)

When making decisions about regrouping strata, the user should also incorporate knowledge about the local HIV epidemic. Consider a scenario in which the "White," "Black," and "Hispanic" race/ethnicity strata meet the 20%/40/10 requirement across all years, but the "Other" race/ethnicity stratum does not meet the requirement for any year being analyzed. The user must then regroup strata in order to combine "Other" with one of the remaining race/ethnicity groups. At this point, the user should rely on what he or she knows about the local HIV epidemic to guide the regrouping decision. For example, if the local epidemic among persons who are not White, Black or Hispanic is known to be most similar to that among Whites, then the user may decide to regroup race/ethnicity strata as "Black," "Hispanic," and "White/Other."

Note: If a particular stratum does not meet the 20%/40/10 requirement but is of particular interest to the local HIS area, then HIV incidence estimation can be performed for that stratum without regrouping. In theory, a stratum can exist as long as it includes at least one BED-recent case. However, the user should keep in mind that the smaller the number of cases in a selected stratum, the more unstable the resulting estimate. For example, using a stratum that contains just one BED-recent case would essentially mean estimating HIV incidence for an entire group of people based on a single case's information. Therefore, CDC does not recommend using any strata that do not satisfy the 20%/40/10 requirement.

The end result of reviewing Table 5 should be a decision as to the desired regrouping of potential strata. The user will perform the actual regrouping in an interactive portion at the beginning of the third SAS program.

### 3.3. PGM3_DETERMINE_STRATA

PGM3_DETERMINE_STRATA begins with an interactive section in which the user should regroup potential strata based on the following instructions:

```
*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*;
*                    ATTENTION: USER INTERACTION REQUIRED!                    *;
*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*;
* Use Table 5 to regroup demographic variables below so that they satisfy the *;
* 20%/40/10 requirement across all selected years.                            *;
*                                                                             *;
* Note: If you are unable to create any strata that satisfy the 20%/40/10     *;
* requirement for a given year or years, then you will only be able to        *;
* generate an unstratified HIV incidence estimate for that year(s).           *;
*                                                                             *;
* Example: You are estimating HIV incidence for years 2006-2008. Your 2006    *;
* data cannot be stratified to meet the 20%/40/10 requirement, but for 2007   *;
* and 2008 you can regroup the agex4 variable to meet the requirement using   *;
* strata for those younger than 30 and those 30 and above. You should regroup *;
* the agex4 variable in the code below as necessary to meet the 20%/40/10     *;
* requirement for both 2007 and 2008, even if the strata will not satisfy the *;
* requirement for 2006--and understand that you will only be able to generate *;
* an unstratified estimate for 2006 (i.e., select 'all' as your 2006          *;
* stratification variable in the next program, PGM4_ESTIMATE_INCIDENCE). You  *;
* will still be able to compare overall estimates across all 3 years, but you *;
```

```
* will only be able to compare stratified estimates for the two age strata    *;
* between 2007 and 2008.                                                      *;
*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*;

%macro strata_group;
length sex2 race4 agex4 mode3 $15;

* Potential stratification variable: Sex (2 default levels)                   *;
if sex = 'M' then sex2 = '1Male';
  else sex2 = '2Female';

* Potential stratification variable: Race/ethnicity (4 default levels)        *;
if race = '6' then race4 = '1White';
  else if race = '4' then race4 = '2Black';
  else if race = '1' then race4 = '3Hispanic';
  else race4 = '4Other';

* Potential stratification variable: Age at infection (4 default levels)      *;
if agex <= 12 then delete;
  else if agex <= 29 then agex4 = '13-29';
  else if agex <= 39 then agex4 = '30-39';
  else if agex <= 49 then agex4 = '40-49';
  else agex4 = '50+';

* Potential stratification variable: Transmission category (3 default levels) *;
if rf = '01' then mode3 = '1MSM';
  else if rf in ('02' '03') then mode3 = '2IDU';
  else mode3 = '3HET/OTH';
%mend strata_group;

*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*;
```

The default potential strata are the two sexes (variable *sex2*), four race/ethnicity groups (*race4*), four age-at-infection groups (*agex4*), and three transmission categories (*mode3*). The age-at-infection variable, *agex4*, is created by adjusting age at diagnosis based on imputed STARHS result and testing group assignments. Since the HIV incidence estimation model applies only to adults and adolescents, all cases that are calculated to have been infected at less than 13 years of age are deleted from further processing. This element of the regrouping code above should not be changed—i.e., pediatric infections should never be included in the HIV incidence estimation process.

Continuing with the earlier example of regrouping transmission category strata (see description of Table 5, page 15), the user may decide to modify the default code as follows in order to meet the 20%/40/10 requirement for 2006, 2007, and 2008:

```
* Potential stratification variable: Transmission category (3 default levels) *;
if rf = '01' then mode3 = '1MSM';
  else mode3 = '2IDU/HET/OTH';
```

However, if the user would rather have the option of estimating HIV incidence among IDU for 2007 and 2008, then he could leave the default strata coding intact:

```
* Potential stratification variable: Transmission category (3 default levels) *;
if rf = '01' then mode3 = '1MSM';
  else if rf in ('02' '03') then mode3 = '2IDU';
  else mode3 = '3HET/OTH';
```

If the user selects the second option, he must remember that *mode3* should not be used as a stratification variable for the 2006 HIV incidence estimate, as the IDU stratum did not satisfy the 20%/40/10 requirement for 2006. Selection of the stratification variable to produce the most robust HIV incidence estimate will be discussed further in section 3.4 below.

After making any necessary changes to the default strata, the user can submit PGM3_DETERMINE_STRATA to generate the final HIV incidence estimation data set, XX_FINAL_HIS_EST_YEAR(S), which includes the newly-defined potential stratification variables. The program also produces the BED-recent heterogeneity review table, which should be used to identify the best stratification variable to be used in estimating HIV incidence for each year.

### *Table 6. Determine Stratification Variable Based on BED-Recent Heterogeneity*

PGM3_DETERMINE_STRATA ends by outputting Table 6 to the **XX Year(s) BED-Recent Heterogeneity Review** PDF file in the "estlib" directory. This table is designed to help the user identify which potential stratification variable—*sex2*, *race4*, *agex4*, or *mode3*—contains the most heterogeneity in percentage of BED-recent cases for each year. Proportions of BED-recent results that differ across strata may indicate that recency of infection differs by stratum, which would further indicate that HIV incidence should be estimated separately for each stratum. Therefore, for each year, the variable with strata encompassing the largest overall range in the percentage of BED-recent cases should be selected as the single "best" stratification variable to produce the most robust HIV incidence estimation. Stratification variables should be selected on a year-by-year basis and can differ from year to year.

Consider the following table as an example:

***XX Year(s) HIV Incidence Estimation Post-Imputation Strata Checks:***
***Table 6. Select Stratification Variable by Year Based on BED-Recent Heterogeneity***

**Year of Diagnosis=YYYY**

| | STARHS Result | | | | |
|---|---|---|---|---|---|
| | 1Recent | | 2Long-Term | | |
| | N | % | N | % | Total |
| **Sex** | | | | | |
| 1Male | 296 | 29.8 | 696 | 70.2 | 992 |
| 2Female | 83 | 25.8 | 237 | 74.2 | 320 |
| **Race/Ethnicity** | | | | | |
| 1White | 140 | 32.4 | 293 | 67.6 | 433 |
| 2Black | 168 | 27.4 | 445 | 72.6 | 613 |
| 3Hispanic | 55 | 23.9 | 177 | 76.1 | 232 |
| 4Other | 14 | 42.2 | 20 | 57.8 | 34 |
| **Age at Infection** | | | | | |
| 13-29 | 146 | 30.8 | 329 | 69.2 | 475 |
| 30-39 | 96 | 26.7 | 264 | 73.3 | 359 |
| 40-49 | 96 | 28.1 | 246 | 71.9 | 342 |
| 50+ | 41 | 29.8 | 95 | 70.2 | 136 |
| **Transmission Category (Imputed)** | | | | | |

| | STARHS Result | | | | |
|---|---|---|---|---|---|
| | 1Recent | | 2Long-Term | | |
| | N | % | N | % | Total |
| 1MSM | 252 | 34.4 | 481 | 65.6 | 733 |
| 2IDU | 41 | 21.2 | 150 | 78.8 | 191 |
| 3HET/OTH | 86 | 22.1 | 302 | 77.9 | 388 |
| | | | | | |
| Total | 378 | 28.8 | 934 | 71.2 | 1,312 |

In this table, race/ethnicity strata encompass the widest range (42.2 – 23.9 = 18.3) in percentage of BED-recent STARHS results for year YYYY. Therefore, race/ethnicity (i.e., SAS variable *race4*) should be selected as the best stratification variable for use in estimating HIV incidence for year YYYY.

Note: All four of the aforementioned potential stratification variables—i.e., *sex2*, *race4*, *agex4*, and *mode3*—are included in Table 6. The strata appear in their regrouped formats, which were determined by the user at the beginning of PGM3_DETERMINE_STRATA. Only the strata that satisfied the 20%/40/10 requirement for a given year are recommended for use in stratifying that year's HIV incidence estimate. To continue with an earlier example, consider the user who really wants to estimate HIV incidence among injection drug users (see description of PGM3_DETERMINE_STRATA, page 17). Assume that the user opted to maintain IDU as a separate stratum, which would enable him to generate IDU-specific HIV incidence estimates for each of the three years. The user should be aware that the 2006 HIV incidence estimate for IDU should not be interpreted (or should be interpreted with extreme caution, recognizing the limitations of the estimate for that stratum). Further, when examining Table 6 output for 2006, the user should ignore the transmission category breakout, as that variable should not be chosen as the stratification variable to produce the most robust 2006 HIV incidence estimate (regardless of any apparent heterogeneity).

### 3.4. PGM4_ESTIMATE_INCIDENCE

PGM4_ESTIMATE_INCIDENCE represents the final step of the HIV incidence estimation process. The program begins with a brief interactive portion in which the user must enter the stratification variable(s) selected during review of Table 6:

```
*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*;
*                    ATTENTION: USER INTERACTION REQUIRED!                    *;
*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*~*;
* Enter list of desired stratification variables (one for each incidence      *;
* estimation year) based on review of Table 6. For each year, among those     *;
* variables that met the 20%/40/10 requirement (either by default or after    *;
* regrouping), select the one with the most heterogeneous distribution of     *;
* percent BED-recent. Stratification variables can differ by year. Enter one  *;
* stratification variable for each year, in order by year, separated by       *;
* spaces. End the list with a semicolon.                                     *;
*                                                                            *;
* Example: You are estimating HIV incidence for 2006 and 2007. All potential  *;
* stratification variables were successfully regrouped to meet the 20%/40/10  *;
* requirement. Sex strata had the most heterogeneity in percent BED-recent    *;
* for Dx year 2006, but transmission category had the most heterogeneity in   *;
* percent BED-recent for Dx year 2007. Set the strata_list variable as:      *;
*                                                                            *;
*         %let strata_list = sex2 mode3                                      *;
```

```
*                                                                          *;
* Note: Selecting "all" for a given year indicates that no available strata   *;
* met the 20%/40/10 requirement for that year and will generate only an       *;
* overall, unstratified estimate.                                          *;
*                                                                          *;
* Example: When estimating HIV incidence for 2006-2008, your available data   *;
* did not meet the 20%/40/10 requirement for any potential stratification     *;
* variable for Dx year 2006. You were, however, able to regroup variables to  *;
* meet the 20%/40/10 requirement for 2007 and 2008. In 2007 the race strata   *;
* showed the most heterogeneity in percent BED-recent, but in 2008 the        *;
* transmission category strata showed the most heterogeneity in percent BED-  *;
* recent. Set the strata_list variable as:                                 *;
*                                                                          *;
*         %let strata_list = all race4 mode3                               *;
*~*~*~
```

| Variable Name | Description |
|---|---|
| *aids* | Total number of concurrent AIDS cases |
| *g0p* | Probability of being detected as BED-recent among new testers |
| *inc* | Total HIV incidence estimate |
| *inc_g0* | HIV incidence estimate for new testers |
| *inc_g1* | HIV incidence estimate for repeat testers |
| *n_g0* | Total number of new testers |
| *n_g0_0* | Number of BED-recent new testers |
| *n_g0_a* | Number of new testers with concurrent AIDS |
| *n_g0_b* | Number of BED-long-term new testers |
| *n_g1* | Total number of repeat testers |
| *n_g1_0* | Number of BED-recent repeat testers |
| *n_g1_a* | Number of repeat testers with concurrent AIDS |
| *n_g1_b* | Number of BED-long-term repeat testers |
| *notest* | Total number of new testers |
| *q* | Percentage of concurrent AIDS among new testers |
| *q1* | Percentage of concurrent AIDS among repeat testers |
| *recent* | Total number of BED-recent cases |
| *reported* | Total number of reported cases |
| *tested* | Total number of repeat testers |
| *v_g0p* | Adjustment for calculating variance among new testers |
| *v_g1p* | Adjustment for calculating variance among repeat testers |
| *v_inc* | Variance of total HIV incidence estimate |
| *v_incg0* | Variance of HIV incidence estimate for new testers |
| *v_incg1* | Variance of HIV incidence estimate for repeat testers |

PGM4_ESTIMATE_INCIDENCE ends by generating a **XX Year HIV Incidence Estimate Results** PDF file for each year included in the estimation. Each file includes overall and stratified HIV incidence estimates (calculated variable *inc*) for a single year, along with the following additional information:

- Number of BED-recent results (*recent*)
- Number of reported diagnoses of HIV disease (*reported*)
- HIV incidence estimate among new testers (*inc_g0*)
- HIV incidence estimate among repeat testers (*inc_g1*)
- Standard deviation of the HIV incidence estimate (*sd*)
- Lower (*low95*) and upper (*high95*) bounds of 95% confidence interval for HIV incidence estimate

Each yearly output file contains HIV incidence estimates for all potential strata included in the regrouping process; however, estimates based on the selected "best," most heterogeneous stratification variable (entered at the start of PGM4_ESTIMATE_INCIDENCE) appear first and should be considered the most robust overall and stratified HIV incidence estimates. Estimates for the other, less-heterogeneous strata can be calculated by applying the stratum-level proportions of total estimated HIV incidence based on the less heterogeneous stratification variable to the overall most robust HIV incidence estimate. Confidence intervals can be similarly calculated by applying an adjustment factor

(overall most robust HIV incidence estimate divided by overall HIV incidence estimate based on less-heterogeneous stratification variable) to the stratum-level confidence intervals based on the less-heterogeneous strata.

For example, consider the HIV incidence estimation process for year YYYY in which the selected "best," most heterogeneous stratification variable was *mode3*. The user would like to calculate HIV incidence estimates by age at infection (*agex4*), too. To do so, the user must refer to output from the *mode3* and *agex4* pages of the **XX YYYY HIV Incidence Estimate Results** PDF file. The first step is to use the "By *agex4*" table to calculate the proportion of total estimated HIV incidence represented by each level of *agex4*:

***By agex4***

| Strata | Number of BED-recent | Number of reported HIV diagnoses | HIV incidence estimate among new testers | HIV incidence estimate among repeat testers | HIV incidence estimate | Standard deviation for HIV incidence estimate | Confidence interval for HIV incidence estimate: lower bound | Confidence interval for HIV incidence estimate: upper bound |
|---|---|---|---|---|---|---|---|---|
| 13-29 | 146 | 558 | 218 | 286 | (0.33) 504 | 108 | 289 | 719 |
| 30-39 | 96 | 459 | 229 | 239 | (0.30) 468 | 106 | 260 | 677 |
| 40-49 | 96 | 412 | 149 | 252 | (0.26) 401 | 90 | 222 | 580 |
| 50+ | 41 | 162 | 73 | 92 | (0.11) 166 | 58 | 51 | 280 |
| | **378** | **1,592** | **670** | **869** | **1,539** | | | |

These proportions are based on an estimate stratified by *agex4*, which was not the most heterogeneous variable and therefore does not represent the most robust estimate. The calculated proportions should then be applied to the overall most robust HIV incidence estimate from the *mode3* output page:

***XX YYYY HIV Incidence Estimate: Stratified by mode3***
***** Estimate Based on Most Heterogeneous Strata *****
***Overall***

| Number of BED-recent | Number of reported HIV diagnoses | HIV incidence estimate among new testers | HIV incidence estimate among repeat testers | HIV incidence estimate | Standard deviation for HIV incidence estimate | Confidence interval for HIV incidence estimate: lower bound | Confidence interval for HIV incidence estimate: upper bound |
|---|---|---|---|---|---|---|---|
| 378 | 1,592 | 643 | 860 | 1,503 | 215 | 1,074 | 1,932 |

The results of applying these proportions are revised stratum-level HIV incidence estimates for *agex4* (based on the more robust *mode3* estimate):

| Strata | Revised HIV Incidence Estimate |
|---|---|
| 13–29 | 0.33 × 1,503 = 496 |

| Strata | Revised HIV Incidence Estimate |
|---|---|
| 30–39 | 0.30 × 1,503 = 451 |
| 40–49 | 0.26 × 1,503 = 391 |
| 50+ | 0.11 × 1,503 = 165 |

Next the user would like to calculate confidence intervals for the newly calculated HIV incidence estimates by age-at-infection strata. The first step is to calculate an adjustment factor, which in this case equals the overall HIV incidence estimate based on *mode3* divided by the overall estimate based on *agex4*: 1,503 ⁄ 1,539 = 0.98. The user should then apply this proportion to the stratum-level confidence intervals in the "By *agex4*" table above to obtain more robust confidence intervals for the age-at-infection HIV incidence estimates:

| Strata | Revised HIV Incidence Estimate | Revised C.I., Lower Bound | Revised C.I., Upper Bound |
|---|---|---|---|
| 13–29 | 496 | 0.98 × 289 = 283 | 0.98 × 719 = 705 |
| 30–39 | 451 | 0.98 × 260 = 255 | 0.98 × 677 = 663 |
| 40–49 | 391 | 0.98 × 222 = 218 | 0.98 × 580 = 568 |
| 50+ | 165 | 0.98 × 51 = 50 | 0.98 × 280 = 274 |

# 4. Interpreting HIV Incidence Estimates

The fifth and sixth CDC-provided SAS programs, PGM5_COMPARE_ESTIMATES and PGM6_EVALUATE_DIFFERENCES, are designed to assist the user in comparing and interpreting one or more HIV incidence estimates.

## 4.1. PGM5_COMPARE_ESTIMATES

PGM5_COMPARE_ESTIMATES should be used when the local HIS areas has generated multiple single-year HIV incidence estimates and should be run immediately following programs 1–4, as it uses the same initial setup completed in PGM1_PREPARE_DATA. Program 5 uses the initial and imputed HIV incidence estimation data sets, XX_HIS_EST_DATA_YEAR(S) and XX_IMPUTED_YEAR(S), to calculate various frequencies and percentages associated with factors that may affect HIV incidence estimates across years. The program ultimately outputs a single table to the PDF file **XX Year(s) HIV Incidence Estimate Comparison** in the "estlib" directory. The table is designed to assist the user in understanding differences that may exist across two or more single-year HIV incidence estimates. More specifically, the table addresses the underlying factors that impact each estimate. All frequencies and percentages displayed in the estimate comparison report are based on the imputed data that were used for the final HIV incidence estimates. Rows shaded in grey are purely informational and are less important than other rows.

Before reviewing the estimate comparison report, the user should take a moment to revisit the foundations of the HIV incidence estimation model. The basic formula is N = R / P, where N is the number of estimated new HIV infections (i.e., HIV incidence), R is the number of BED-recent STARHS results, and P is the estimated probability of being detected as BED-recent at HIV diagnosis. In a simplistic scenario in which P is held constant, an increase in R will lead to a higher HIV incidence estimate, and a decrease in R will lead to a lower HIV incidence estimate. Because HIV incidence is estimated separately for new and repeat testers and summed to generate a final estimate, the formula can also be written as $N = N_{new} + N_{rep} = (R_{new} / P_{new}) + (R_{rep} / P_{rep})$.

By definition, new testers are tested less frequently than repeat testers and thus have a lower likelihood of being detected as BED-recent at HIV diagnosis. Therefore, each BED-recent new tester carries more weight in the estimation process than each BED-recent repeat tester. Furthermore, because new and repeat testers have different testing behaviors, their probabilities of being detected as BED-recent at HIV diagnosis are estimated differently:

- The most important parameter in estimating the new tester's probability of being detected as BED-recent at HIV diagnosis ($P_{new}$) is the proportion of concurrent HIV/AIDS diagnoses (q). $P_{new}$ is inversely related to q: The higher the proportion of concurrent HIV/AIDS diagnoses, the lower the probability of a new tester being detected as BED-recent at HIV diagnosis. If the number of BED-recent cases ($R_{new}$) is held constant, then the HIV incidence estimate for new testers ($N_{new}$) will increase when q increases.
- The most important parameter in estimating the repeat tester's probability of being detected BED-recent at HIV diagnosis ($P_{rep}$) is the time interval between the last negative test and the first positive test (T). $P_{rep}$ and T are inversely proportional as well: The longer the time interval between last negative and first positive tests, the smaller the probability that a repeat tester will be detected as BED-recent at HIV diagnosis. If the number of BED-recent cases ($R_{rep}$) is held constant, then the HIV incidence estimate for repeat testers ($N_{rep}$) will increase when T increases.

Two factors that are related the numbers of BED-recent results among both new and repeat testers ($R_{new}$ and $R_{rep}$, respectively) are:

- Number of new HIV diagnoses: Typically, an increase in true HIV incidence for a given year will result in a commensurate increase in HIV diagnoses; therefore, the more new diagnoses of HIV disease occur in a given year, the higher the estimated HIV incidence for that year will be.
- Proportion of BED-recent: The proportion of BED-recent results determines the number of observed recent infections. Typically, the higher the proportion of BED-recent results, the higher the HIV incidence estimate will be.

In the absence of evidence suggesting changes in testing behavior (i.e., changes in q and/or T), a change in the proportion of BED-recent results may indicate a real change in HIV incidence. On the other hand, if testing behavior does change suddenly, then an accompanying change in the HIV incidence estimate does not necessarily represent a true change in HIV incidence. One of the underlying assumptions of the

estimation model is that testing behavior has not changed in recent years; therefore, the model is sensitive to sudden changes in testing, and a change in estimated HIV incidence that is accompanied by sudden changes in testing may not represent a change in true HIV incidence.

With this understanding of how the number of new HIV diagnoses, proportion of concurrent AIDS diagnoses (q), intertest interval (T), and proportion of BED-recent results interact, the user can examine the **XX Year(s) HIV Incidence Estimate Comparison** PDF output to evaluate differences across two or more single-year HIV incidence estimates. The comparison report should be reviewed along with the HIV incidence estimate results tables generated by PGM4_ESTIMATE_INCIDENCE. Consider the sample 2006–2008 HIV incidence estimates for area XX and the accompanying HIV incidence estimate comparison report:

| Year | HIV Incidence Estimates | | |
|---|---|---|---|
| | New Testers | Repeat Testers | Total |
| 2006 | 893 | 686 | 1,579 |
| 2007 | 643 | 860 | 1,503 |
| 2008 | 1,024 | 679 | 1,703 |

***XX 2006-2008 HIV Incidence Estimate Comparison:***
***Comparing Potential Factors Affecting HIV Incidence Estimates Across Years***

| | 2006 | | 2007 | | 2008 | |
|---|---|---|---|---|---|---|
| **Factor** | **N** | **%** | **N** | **%** | **N** | **%** |
| All new HIV diagnoses | 1,600 | . | 1,600 | . | 1,600 | . |
| New testers | 648 | 40.5 | 685 | 42.8 | 659 | 41.2 |
| AIDS at HIV Dx | 331 | 20.7 | 285 | 17.8 | 295 | 18.4 |
| AIDS Dx 1-6 mos. after HIV | 176 | 11.0 | 196 | 12.3 | 187 | 11.7 |
| With STARHS results | 1,093 | 68.3 | 1,119 | 69.9 | 1,118 | 69.9 |
| Among cases with STARHS results, BED-recent | 342 | 31.3 | 378 | 33.8 | 370 | 33.1 |
| **New testers** | | | | | | |
| Among new testers with STARHS results, BED-recent | 89 | 24.0 | 79 | 19.8 | 111 | 27.7 |
| Among new testers, AIDS at HIV Dx | 177 | 27.3 | 177 | 25.8 | 172 | 26.1 |
| **Repeat testers** | | | | | | |
| Among repeat testers with STARHS results, BED-recent | 252 | 34.9 | 299 | 41.6 | 260 | 36.3 |
| Among repeat testers, AIDS at HIV Dx | 154 | 16.2 | 108 | 11.8 | 123 | 13.1 |
| Among repeat testers, average T (in months) | 32 | . | 39 | . | 33 | . |
| Among repeat testers, median T (in months) | 13 | . | 24 | . | 24 | . |

A few points of interest:

- The number of new diagnoses of HIV disease stayed constant across all three years. The percentage of new testers also stayed roughly the same.

- The proportion of concurrent AIDS diagnoses was slightly higher in 2006, but the number of BED-recent results was lower in 2006 than in other years; therefore, the proportion of concurrent AIDS diagnoses should not have a big impact on the HIV incidence estimates.
- The overall proportion of BED-recent results was about the same across all three years, but it differed quite a lot between new and repeat testers. The 2008 percentage of BED-recent results among new testers was higher than in 2006 and 2007, and this explains the similar increase in the 2008 HIV incidence estimate among new testers.
- Because proportions of concurrent AIDS diagnoses did not change much across the years, testing behavior should not have changed much, either; however, the 2008 percentage of BED-recent results among new testers increased, so the higher 2008 HIV incidence estimate among new testers may represent a true increase in HIV incidence.
- The 2007 percentage of BED-recent results among repeat testers was much higher than the same percentage for either 2006 or 2008. The average T for repeat testers in 2007 was longer than in 2006 and 2008, and the percentage of new testers was slightly higher in 2007, suggesting either that some of the newly-diagnosed repeat testers had not been tested in quite awhile or that repeat testers tested less frequently. The cumulative effect of increased percentage of BED-recent results and longer inter-test interval (T) resulted in an increased 2007 HIV incidence estimate among repeat testers. However, because the data indicate that there might have been a change in testing behavior, it is difficult to determine whether the increased estimate represents a true increase in HIV incidence for repeat testers in 2007.

## 4.2. PGM6_EVALUATE DIFFERENCES

PGM6_EVALUATE_DIFFERENCES provides additional tools to assist the user in interpreting local HIV incidence estimates. This program requires the user to enter various numbers by hand but does not refer to any input data sets; therefore, PGM6_EVALUATE_DIFFERENCES can be run at any time. The sixth and final program is designed to allow for future additions; at present, the program allows the user to (1) calculate confidence intervals for HIV incidence rate estimates and (2) determine whether two HIV incidence estimates are significantly different. The program requires user input as described in the following instructions:

```
*****************************************************************************;
* 1. To calculate confidence intervals for HIV incidence rate estimates,      *;
* values for the following macro variables: inc (HIV incidence estimate),     *;
* se_inc (standard error of HIV incidence estimate), pop (population          *;
* estimate), and se_pop (standard error of population estimate).             *;
*****************************************************************************;
%let inc = ;
%let se_inc = ;
%let pop = ;
%let se_pop = ;

*****************************************************************************;
* 2. To determine whether two HIV incidence estimates are significantly       *;
* different, enter values for inc1 (HIV incidence estimate #1), se1 (standard *;
* error of HIV incidence estimate #1), inc2 (HIV incidence estimate #2), and  *;
* se2 (standard error of HIV incidence estimate #2).                         *;
*****************************************************************************;
%let inc1 = ;
```

```
%let se1 = ;
%let inc2 = ;
%let se2 = ;
```

Results of both calculations are printed to the SAS Output window.

## 5. Contact Information

If the user has any questions or concerns about this document or local HIV incidence estimation in general, he or she should contact the HICSB epidemiologist assigned to provide technical assistance to his or her HIS area. The main HICSB telephone number is (404) 639-2050.

# Appendix 7- Documentation for HIV Incidence Completeness Report

## Purpose

The HIV incidence completeness report provides a snapshot of the proportion of HIV cases diagnosed within a three year period that have required TTH data elements and BED results. The HIV incidence completeness report should serve as a data quality monitoring tool for CDC epidemiologists and state/local coordinators to measure HIS data completeness, identify problems with data transfer and monitor progress toward meeting target performance levels and data quality standards. Even though the criteria used in the HIV incidence completeness report are not necessarily the criteria used for the annual national evaluation, high completeness rates will support the HIV incidence surveillance programs to meet the outcome standards. If issues are identified CDC will provide technical assistance to improve the completeness of the data. The HIV incidence completeness report will be generated on a monthly basis starting in January 2010.

## Description

The HIV incidence completeness report is composed of three sections: (1) completeness of TTH required data elements; (2) completeness of cases having a sample with STARHSID and BED result; and (3) proportion of data contributing unique information for HIV incidence analysis.

## Section 1: Completeness of TTH required data elements

This section reflects an HIV incidence surveillance program's efforts in collecting data for the following TTH data elements:

- TTH date of first positive HIV test
- Ever tested HIV negative
- Number of negative tests in the 2 years before first positive (including the first positive – following eHARS implementation only negative test before first positive tests will be included, the first positive test will not be included)
- Ever taken ARVs or ARV use dates
- TTH date of last negative HIV test
- ARV use dates
- Completeness of TTH

Completeness is defined as the value for the data element is not blank; the field is deemed complete as long as there is a response to the variable, including responses of "Refused" or "Unknown". For date variables, the value can be 9999, as long as incidence surveillance uses the IVR database. After the transition to eHARS, that value will not be available. Completeness of each TTH data element is calculated separately.

A complete TTH is defined as having a value for least one of the six TTH data elements. The six data elements are: (1) TTH date of first positive; (2) ever tested HIV negative; (3) number of negative tests in the 2 years before first positive; (4) TTH date of last negative HIV test; (5) ever taken ARVs; (6) ARV use dates..

The completeness of TTH date of first positive, ever tested negative, number of negative tests in the past 2 years before first positive, ever taken ARVs or dates of ARV use, and completeness of TTH are based on all HIV cases diagnosed in a specific year. The completeness of TTH date of last negative is

based on cases that have a "Yes" for ever tested negative. The completeness of ARV dates is based on cases that have a "Yes" for ARV use.

### Section 2: Completeness of cases having a sample with STARHSID and BED result

This section measures the completeness of cases

- With a STARHSID
- Without a STARHSID but having a reason for not testing
- With BED result of recent or long term

All measures are based on all HIV cases diagnosed in a specific year.

The measure of cases with a STARHSID determines the proportion of cases having a sample that has been identified by the incidence program as "to be tested" from a test/toss list or by the state's public health laboratory. The second measure (cases without STARHSID but having a reason for not testing) determines the proportion of cases that do not have a sample for testing, but a reason for not testing has been identified by the incidence program. The third measure (cases with BED result) determines the proportion of cases having a sample with a BED test result either "Recent" or "Long term". Cases with a BED result are a subset of cases having a sample with STARHSID. Therefore, from the difference of the two measures (cases with a STARHSID and cases with BED result), sites may identify cases that have pending samples for BED results or did not have a test result. Significant differences between the two measures need to be investigated.

### Section 3: Proportion of data contributing unique information for HIV incidence analysis

This section measures the proportion of data that contributes unique information for HIV incidence estimation. Data contributing unique information for analysis include responses that help to classify a case as a repeat or new tester or having a recent vs. long standing infection that would only be obtained as a part of HIV incidence surveillance.

The proportion of TTH data contributing unique information for analysis is based on all HIV cases for a specific year. TTH data contributing unique information for analysis is defined as having at least one valid answer for any of the six required data elements:

1. TTH date of first positive HIV test has a valid year and a date that is earlier than the eHARS diagnosis date.
   Note: For HIV incidence analyses, the eHARS diagnosis date is selected when TTH date of first positive is the same as eHARS diagnosis date. However additional information to re-classify a case as recent or long term infection is available if the TTH date of first positive is earlier than the eHARS diagnosis date. If the TTH date of first positive is 6 months or earlier than the eHARS diagnosis date, the BED results of the case will be set to "Long term". If the TTH date of first positive is 1-6 months earlier than eHARS diagnosis date, the BED results will be set to missing.
2. Answer to ever tested HIV negative is "Yes" or "No".
   Note: If the answer to ever tested negative is "Yes", the case is classified as a repeat tester; if the answer is "No", the case is classified as a new tester. Surveillance staff should never enter "No" if previous testing status is not known.
3. Answer to number of HIV negative tests in the 2 years before first positive is not blank, "Refused" or "Don't know".

Note: If the number of HIV negative tests in the 2 years before first positive is >=2, the case is classified as a repeat tester, and if the answer is "0", the case is classified as a new tester. Currently CDC is unable to use the value "1" because it has been inconsistently collected.

4. Date of last negative test year must have a valid year.
   Note: If there is a valid date of last negative test, the case is classified as a repeat tester.
5. Answer to ever taken ARVs is "Yes" or "No"
6. Dates of ARV use exist and are valid.
   Note: If answer to ever taken ARVs is "Yes" and if the ARV use is during the six months prior to collection of the specimen used for the BED test, the BED result of the case is set to missing.

The proportion of BED data contributing unique information for analysis is defined as having BED test results of "Recent" or "Long term". The measure is calculated in two ways. The first measure excludes cases with AIDS diagnosed within 6 months of HIV diagnosis from the denominator. In the analysis, the BED result for cases with AIDS diagnosed within 1- 6 months of HIV diagnosis is set to "Long term", regardless of their original BED result. This measure is useful for assessing the completeness for cases diagnosed in the previous years, but because it does not allow time for reporting delay, it is not as helpful in assessing completeness in the current year.

The second measure of BED completeness is useful to assess cases diagnosed in the current year. In the second measure, cases with AIDS diagnosed in the same month of HIV diagnosis are excluded from both numerator and denominator.

- **How to use the HIV incidence completeness report:**

1. The denominator includes cases reported from each HIS area and residing in areas funded for HIS.

2. The proportion of data contributing unique information to analysis is added to emphasize the content of the data rather than the collection and reporting of the data.

3. The completeness of each data element and completeness of TTH (Section1) reflects the overall efforts in data collection; all values (as long as there is a response in the field, including responses of "Refused" or "Unknown") are accepted. The proportion of TTH data contributing unique information for incidence analyses (Section 3) only counts answers that provide unique information for analysis, therefore, it is usually lower than the completeness proportion of TTH or each data element. Among the TTH data elements, completeness of TTH date of first positive is particularly more than the TTH date of first positive contributing unique information for incidence analysis. As the eHARS diagnosis date variable is used for incidence analyses, the TTH date of first positive only provides unique information for analysis if it has a valid date that is earlier than the eHARS diagnosis date. It is important to highlight that TTH variable values of "Refused" or "Unknown" do not provide additional information for analysis.

4. The proportion of BED results contributing unique information for incidence analysis for cases diagnosed in the previous years is higher than for cases diagnosed in the current year due to delays in reporting time.

# Documentation for HIV Incidence Error Report

## Purpose

The HIV incidence error report is a tool to monitor data entry errors and inconsistencies between variables and improve the quality of HIV incidence surveillance (HIS) data. The purpose is to support the efforts of state/local incidence coordinators in improving the validity and accuracy of the data and monitoring progress toward meeting target performance levels and data quality standards. If errors are checked and fixed routinely, HIS professionals can detect problems with data entry and data collection, and improve data quality. The HIV incidence error report is generated quarterly (June, September, December, and March). Errors should be resolved prior to submission of incidence datasets in August, November, February and May.

## Description

The HIV incidence error report is an excel file titled XX_error_yyyymm.xls (where xx is the 2-letter or 3-letter as appropriate abbreviation for the state/local incidence area) with an error summary (worksheet 1) and an error details (worksheet 2).

Worksheet 1: Error Summary

The error summary is a summary report of all the errors, including types and the frequencies of each error. The error summary has three columns: Errormsg, Errortype and Count. Errormsg is the description of each error. Errortype identifies whether the error is related to TTH, laboratory or core surveillance, and Count is the column for frequency, or number of records with that error (see example below). Based on the frequency of the errors, program areas can identify the most frequently occurring errors and prioritize corrective actions.

| *Errormsg* | *Errortype* | *Count* |
|---|---|---|
| 3. TTH date of last neg after eHARS date of diagnosis of HIV infection | TTH | 10 |

Worksheet 2: Error Details

The error details sheet provides case-level error information. Errors are sorted by diagnosis year, stateno/cityno, and entered year. The error detail worksheet has 11 columns:

| | **Column Name** | **Column Description** |
|---|---|---|
| 1 | PID | A Pseudo ID variable was created in order to link the error report and the error resolution file records together. This is important because the State_city_No column contents are cleared from the resolution file before the files is sent to CDC. |
| 2 | State_city_No | State_city_No is a unique identifier that helps in the identification of the case. |
| 3 | Errortype | Determines whether the error is related to TTH or laboratory. |
| 4 | Entered_yr | Provides the year that the case was entered into eHARS |
| 5 | Dx_yr | Provides the year the case was diagnosed as HIV infection that is the earlier date between HIV diagnosis date and AIDS diagnosis date. |
| 6 | CheckingID | Helps with the identification of the case using document_uid or starhsid. HIS data in eHARS is document based with multiple TTH and Lab documents. Therefore, this new variable has been added to the error report to provide additional help in identifying which document within a case is potentially problematic. |
| 7 | VariableA | Displays the variable names and values with errors. VariableA is used in combination with VariableB to identify data entry errors and inconsistencies between variables. |
| 8 | VariableB | Displays variable names and values with errors. VariableB is used in combination with VariableA to identify data entry errors and inconsistencies between variables. |
| 9 | Errormsg | Provides a description of the error. The table below displays each error message. |
| 10 | Fixable | Sites should record whether the error can be resolved or not using the values “YES” or “NO”. |
| 11 | Comment | Sites should indicate the reason the error is not fixable using the codes A-E or specify any other comments described below. |

The first 9 columns (PID, State or city no, Errortype, Entered_yr, Dx_yr, CheckingID, Variable A, Variable B and Errormsg) are completed by CDC.
The columns Fixable and Comment are for program areas to use. Fixable is used to record whether the error can be resolved or not; values for Fixable field should be "YES" or "NO". If the error is fixable the program area should indicate "YES" in the Fixable field and correct the error in eHARS; fixed errors will be reflected in the monthly SAS datasets and will therefore not reappear in subsequent error reports. If the error is not fixable (the data must remain as entered) the program area should indicate "NO" in the Fixable field and provide additional comments in the Comment field. Likewise, the data inconsistencies will also not reappear on subsequent error reports. Use the list below to enter the code for the most commonly used comments:

A. Self-reported information:
Data entered correspond to the original data collection form provided by the patient, provider or DIS.

B. Unable to determine from chart abstraction
DIS or incidence staff re abstracted data from chart, but they were not able to fix the error based on the information obtained.

C. Unable to follow up
DIS or incidence staff was unable to conduct further investigation to fix the error. Some possible reasons include that the patient was lost to follow up (e.g. diagnosed previously in another state, or patient is unable to be re interviewed), or the chart was not available for re-abstraction.

D. Ineligible client
Patient is not a new diagnosis or is from a jurisdiction not participating in HIS.

E. Data verified
This is specifically for specimens with an obtained date that is not within 0-3 months of eHARS date of diagnosis of HIV infection. Due to the nature of the data collection and reporting process available specimens may not have been drawn within 0-3 months of the eHARS date of diagnosis of HIV infection and program areas can do nothing about it. For example, a specimen was sent for STARHS and an earlier diagnosis date was subsequently found.

If there are other comments beyond this list, please briefly describe in the Comment field.

The error details worksheet is protected. As a result, program areas cannot delete rows or columns or make changes to the worksheet except for four columns: State_city_no, CheckingID, Fixable and Comment. Fixable and Comment are used to record resolutions for errors and the content of State_city_no and CheckingID columns must be cleared before the resolution file is sent to CDC.

**Data Error Checks and Recommended actions**

***General comments:***
Data errors are errors associated with data entry or inconsistencies between variables. There are three categories of errors: TTH-related, laboratory-related and core surveillance related. Based on the categories of the errors, program areas can refer to relevant documents or personnel to investigate. All errors identified in the error report should be investigated by verifying the source documents or conducting additional follow-up activities. Data entry errors should be corrected after investigation.

Data inconsistencies reflect problems associated with data collection and reporting process that should be investigated. Inconsistencies in the data may require additional follow up activities including re-abstraction of data, contacting the reporter for clarification or re-interview of the patient if a short time has elapsed between the original interview and the time the error is identified. If systematic errors or inconsistencies are identified re-training of data collection staff and reporters may be necessary. Training may need to include review of interviewing skills and data abstraction procedures. CDC has developed Job Aids to assist all HIS sites in conducting these training activities.

The following table contains error checks, action priority and suggestions for how they may be addressed. Astute observers will note that two errors on the previous error report are no longer reported.

- Error #6 has been deleted and replaced by error #5 which is a combination of the former Errors #5 and #6 from the previous error report.

- Error #8 “BED result exists, specimen collection date before eHARS date of diagnosis of HIV infection” has been deleted in order to avoid redundancy as Error 8 is a subset of Error #13.

| Priority | Error (errormsg) | Type | Comments |
|---|---|---|---|
| 1 | BED result exists, but STARHSID missing | Lab | Check all STARHS laboratory reports related to the case to obtain the STARHSID. |
| 2 | BED result exists, but specimen collection date missing | Lab | Check all laboratory reports to obtain the specimen collection date. |
| 3 | TTH date of last negative after eHARS date of diagnosis of HIV infection | TTH | Refer to original source of report for the TTH and laboratory reports to verify. |
| 4 | TTH date of last negative after TTH date of first positive | TTH | Refer to original source of report for the TTH and laboratory reports to verify. |
| 5 | If “ever tested negative” is “refused”, “unknown” or blank but number of negative tests >=1. | TTH | Verify TTH source for data entry errors. If not data entry errors, further investigation may be required. If a systematic error is detected re-training of the staff and reporters may be required. |
| 6 | Blank | Blank | Error #6 has been deleted and replaced with Error #5 |
| 7 | BED result exists, specimen collection date after STARHS test date | Lab | Check laboratory related reports and STARHS result excel files. |
| 8 | Blank | Blank | Error #8 been deleted as it is a subset of Error #13. |
| 9 | TTH date of first positive before 1975 | TTH | Check for data entry or record error. |
| 10 | TTH date of first positive before birth date from eHARS | TTH | Verify birth date and verify TTH date of first positive. |
| 11 | If TTH last negative date within 24 months of TTH first positive date, and “number of negative tests* within 24 months before first positive“ < 1 | TTH | Verify TTH source for data entry errors. If not data entry errors, further investigation may be required. If a systematic error is detected re-training of the staff and reporters may be required. |
| 12 | If “ever tested negative” is “no”, “refused”, “unknown” or blank, and “date of last negative test “ exists | TTH | Verify TTH source for data entry errors. If not data entry errors, further investigation may be required. If a systematic error is detected re-training of the staff and reporters may be required. |
| 13 | BED result exists, specimen collection date not within 0-3 months of eHARS date of diagnosis of HIV infection* | Lab | Check the accuracy of the specimen collection date. This error is very likely due to the availability of the specimen and the nature of data collection and reporting process. |
| 14 | Ever tested positive is Yes, first positive test date is missing | TTH | Verify TTH source for data entry errors. If not data entry errors, further investigation may be required. If a systematic error is detected re-training of the staff and reporters may be required. |
| 15 | First positive date after eHARS date of HIV diagnosis | TTH | Verify TTH source for data entry errors. If not data entry errors, further investigation may be required. If a systematic error is detected re-training of the staff and reporters may be required. |
| 16 | Invalid self-reported last negative date | TTH | Verify TTH source for data entry errors. If not data entry errors, further investigation may be required. If a systematic error is detected re-training of the staff and reporters may be required. |
| 17 | Invalid specimen collection date | Lab | Check the accuracy of the specimen collection date. |
| 18 | Invalid SOD values | Lab | Check for data entry or record error. |

***How to use the error report:***
The error report is generated quarterly at the beginning of June, September, December and March. Error Reports are uploaded to the SDN for areas to retrieve with a password of XXDQ (where XX represents the site name abbreviation). CDC epidemiologists will discuss progress on resolving the incidence error report as a part of the routine monthly call.

Project areas will investigate and determine the resolutions for errors.
(1) Fixable errors should be corrected in eHARS and the corrections will be reflected in monthly SAS datasets.
(2) For errors that are not fixable, record "No" in the Fixable column, and provide reasons in the Comment column using the comment list provided. Any error marked "No" will not appear on future error reports.
(3) When finished, clear the content of the State_City_No and CheckingID columns in the Error_details worksheet, save the file as XX_error_yyyymm_resolution.xls (where XX is the project area abbreviation) and email it to CDC incidence data manager (Mona Doshani – MDoshani@cdc.gov or imf6@cdc.gov) by Aug 15, Nov15, Feb15 and May15.

Because the error details worksheet is protected, columns cannot be deleted, but the contents of the columns can be cleared. To clear the content of the State_city_no and the CheckingID columns prior to returning the resolution file to CDC, right click on the State_city_no and the CheckingID column letter and select Clear Contents from the popup menu. The resolution file must be sent back to the CDC incidence data manager by August 15, November 15, February 15 and May 15, because the resolutions of the errors will be incorporated in the subsequent error reports. Errors that have been fixed or investigated but with no possible resolution will not appear in the subsequent error reports.

It is recommended that HIS program areas keep a copy of the resolution file with State_city_no and CheckingID in case consultation with the CDC epidemiologist or data manager is needed regarding certain data problems.

Note: Following the *NCHHSTP Data Security and Confidentiality Guidelines for HIV, Viral Hepatitis, STD, TB programs: Standards to Facilitate Sharing and Use of Surveillance Data for Public Health Action*, absolutely NO State_city_no and CheckingID column should be sent in emails. Sending confidential data through email will result in having to use SDN to send the resolution file back to CDC in the future.