WOMEN'S INTERAGENCY HIV STUDY
SECTION 20: REGISTRY MATCH PROTOCOLS

A. OBJECTIVES

All WIHS sites will perform registry matches at specified intervals (see B below) to:

1. Determine the incidence of death, tuberculosis (TB), cancer, and AIDS among WIHS participants; and
2. Measure the validity of self-reported medical information.

B. REGISTRIES SEARCHED

- National Death Index Plus (NDI): performed yearly, after the new NDI plus data is available
- County Vital Statistics: performed as needed by the sites
- County/State Tuberculosis Control Registry: performed as specified by the EC
- State Cancer/Tumor Registry: performed every two years or as specified by the EC
- County AIDS Surveillance (optional)

C. CRITERIA FOR SUBMITTING MATCHES

1. WHO TO SUBMIT

To ensure uniformity of submissions across the clinical sites, each site should use the same criteria for determining whom to submit.

a. Death Registries

All recruits who are not actively being followed, excluding those women who specifically asked to be disenrolled and have their names and records destroyed. **Do not submit only the names of those known to be deceased – this will create ascertainment bias.** Even if the site has a copy of the death certificate, the participant should be submitted to the NDI Plus to get the electronically coded causes of death information. Once this electronic information is received, that individual does not need to be submitted to NDI Plus again.

b. TB and Cancer Registries

All recruits who have signed a consent form allowing their names to be submitted for TB and Cancer registry searches, excluding those women who specifically asked to be disenrolled and have their names and records destroyed. Since TB and cancer can occur more than once during the life of the individual, all eligible participants should be searched every time these registry matches are performed. Women who died or were lost to follow-up during WIHS I and thus have not signed a consent form can be searched for if the following applies:

> According to Section 116D of the OPRR (Office for Protection from Research Risks) regulations concerning human subjects, an IRB can waive consent under the following three conditions: 1) minimal risk to the person, 2) no violation of the person's rights or welfare, and 3) no practical way to get reconsent.

Both the WIHS Executive Committee and National Community Advisory Board have approved allowing searches for those women who have not signed a consent form and who meet these three conditions. Sites will need to receive additional approval from their local IRBs.

c. AIDS Registries

Who to match is based on site-specific consent and medical record release forms.

NOTE: Each time a list of WIHSIDS is submitted to a registry for matching, this list should be sent to calden@jhsph.edu for archiving. Data Managers should also keep this data locally.

2. WHAT REGIONS TO SEARCH

For national death data, search the NDI. For TB and cancer data, sites should search the state(s) registry where the majority of their participants reside. If state TB registries are not appropriate (as is the case in NYC), then local searches should be done. For local death and AIDS data, sites should search those counties where the majority of their participants reside. Be sure to include participants who have transferred in or out of your study site, as they may have had a TB, AIDS or cancer diagnosis while a resident in your state.

3. FREQUENCY AND TIMING OF SUBMITTALS

NDI Plus data is updated each year and includes cause of death information for deaths that have occurred in the year two years prior to the current year. The WIHS EC currently specifies that sites submit NDI requests every year in order to obtain the most recent data available.

Local death registry (vital statistics) matches should be done annually, as soon as data from the previous year is available (usually around February).

The WIHS EC currently specifies that sites submit to state Cancer registries every other year (even-numbered years) to obtain the most recent data available.

The WIHS EC will decide on a year-to-year basis if sites should submit to state or local TB registries. The most recent TB match was done in 2007.

Each site should find out from their state registries what time period was covered by their last TB and Cancer matches and give this information to WDMAC. In addition, sites need to give WDMAC the list of WIHSID numbers submitted to each registry (Cancer, TB, AIDS) to determine the denominator for person-years of observation.

AIDS registry matches are performed on an ad hoc basis, as determined by site specifications.

D. PROCEDURES FOR SUBMITTAL

1. NATIONAL DEATH INDEX PLUS (NDI)

Download application materials from the NDI web site:
http://www.cdc.gov/nchs/data/ndi/ndi_application.pdf.

Complete the application, submission and resubmission forms as described for the NDI Plus. Submit data sets for matching as described in the NDI users' manual. Allow two to four weeks for processing of submitted materials.

2. TB REGISTRY MATCHES

- Several WIHS sites have study participants from more than one county. Therefore, the most efficient way to match is on the state level. The first step is to figure out if your state has a TB registry. This can be done by using an internet search engine and typing in the name of your state and the word TB registry. If matching on the state level is not possible or appropriate (as is the case for NYC), you will need to determine in which counties the majority of your site's participants reside. Virtually all large cities have a TB control program. Find out who is in charge of each and explain to those people what you want to do.
- Write a letter of introduction that briefly explains who you are and what you would like to do. Enclose a copy of your current human subjects approval letter, an abbreviated WIHS protocol

(e.g., the one prepared each year for your IRB), and a draft protocol for performing the match (see Appendix A to this section). To save time and attract immediate attention, use overnight mail for transmitting these materials to TB control.

- Follow up with a phone call to find out if there are any problems with gaining access to the registry data. You will also need to work out the logistics with TB control, especially whether the match will be done by computer or by hand. All of the data requested is on a computerized form that goes to the CDC, but that doesn't mean that each TB control program will be able to perform a computer match. It does mean, however, that all the data are available. If the data need to be retrieved by hand, perhaps your site's medical record abstractionist can perform the task.

3. AIDS REGISTRY MATCHES

   These are done on the local level only. Follow the above procedures for the TB registry match.

4. CANCER REGISTRY MATCH

   Contact information for U.S. State Cancer/Tumor registries can be found by searching the web. Contact the designated individual at the State Cancer registry and find out what the process is for getting approval and completing a computerized Cancer registry match. Procedures and costs of the computerized match vary widely by state.

   Some useful web links for contacting Cancer registries include:

   - North American Association of Central Cancer Registries (NAACCR) – http://www.naaccr.org/ – NAACCR is a collaborative umbrella organization for Cancer registries, governmental agencies, professional associations, and private groups in North America interested in enhancing the quality and use of Cancer registry data. Most central Cancer registries in the United States and Canada are members.
   - National Program of Cancer Registries – www.cdc.gov/NCCdphp/dcpc/npcr/index.htm
   - Surveillance Epidemiology and End Results (SEER) Program – www.seer.ims.nci.nih.gov/
   - California Cancer Registry – www.ccrcal.org
   - New York State Department of Health – www.health.state.ny.us/nysdoh/cancer/cancer.htm
   - Illinois Department of Public Health, Epidemiology and Health Systems Development – www.idph.state.il.us/about/epi/index.htm

E. REPORTING RESULTS

1. NATIONAL DEATH INDEX AND LOCAL DEATH REGISTRY MATCHES

   As per the Outcomes Ascertainment Protocol (Manual of Operations, Section 12), a *Disenrollment Form (DENR)* must be completed and entered into Apollo for every death discovered (through passive surveillance or registry match). Data entry of the DENR will cause an *Ascertainment Tracking Checklist (ATC)* to be generated, which will trigger site staff to transmit cause of death information (via the *CORE Form*) to WDMAC for inclusion in the OUTCOME.DAT summary file.

   Since outcomes data will be analyzed by events falling within a specific time interval versus within a specific visit, when the first report of a death is through registry match, the visit number for the death should be assigned based on the following time intervals:

10/1/1994	3/31/1995	1	
4/1/1995	9/30/1995	2	
10/1/1995	3/31/1996	3	
4/1/1996	9/30/1996	4	
10/1/1996	3/31/1997	5	
4/1/1997	9/30/1997	6	WIHS I
10/1/1997	3/31/1998	7	
4/1/1998	9/30/1998	8	
10/1/1998	3/31/1999	9	
4/1/1999	9/30/1999	10	
10/1/1999	3/31/2000	11	
4/1/2000	9/30/2000	12	
10/1/2000	3/31/2001	13	
4/1/2001	9/30/2001	14	
10/1/2001	3/31/2002	15	
4/1/2002	9/30/2002	16	WIHS II
10/1/2002	3/31/2003	17	
4/1/2003	9/30/2003	18	
10/1/2003	3/31/2004	19	
4/1/2004	9/30/2004	20	
10/1/2004	3/31/2005	21	
4/1/2005	9/30/2005	22	
10/1/2005	3/31/2006	23	
4/1/2006	9/30/2006	24	
10/1/2006	3/31/2007	25	WIHS
4/1/2007	9/30/2007	26	III
10/1/2007	3/31/2008	27	
4/1/2008	9/30/2008	28	
10/1/2008	3/31/2009	29	
4/1/2009	9/30/2009	30	
10/1/2009	3/31/2010	31	
4/1/2010	9/30/2010	32	
10/1/2010	3/31/2011	33	
4/1/2011	9/30/2011	34	
10/1/2011	3/31/2012	35	WIHS
4/1/2012	9/30/2012	36	IV
10/1/2012	3/31/2013	37	
4/1/2013	9/30/2013	38	
10/1/2013	3/31/2014	39	
4/1/2014	9/30/2014	40	
10/1/2014	3/31/2015	41	
4/1/2015	9/30/2015	42	
10/1/2015	3/31/2016	43	
4/1/2016	9/30/2016	44	
10/1/2016	3/31/2017	45	
4/1/2017	9/30/2017	46	WIHS V

The actual visit number during which a death occurred should be reported on the *DENR*, not the visit number during which the death was discovered. The *DENR* should be updated with the new visit number if additional information regarding the date of death is found.

However, if a participant completed a visit and then died during the same visit window, visit number should be recorded on the *DENR* as the visit number of the next visit window. This will prevent data from being overwritten in the VISITS table of Apollo.

a. National Death Index Plus

Approximately two to four weeks after submitting to the NDI, sites should expect to receive an electronic file (File #9) containing the NDI cause of death information. Sites should then complete the *DENR* as described above for each reported match. Entry of the *DENR* will generate an *ATC* form. Death certificates need not be requested for participant deaths as the NDI Plus matching will provide cause of death information.

After completion of the *DENR*, sites should transmit the NDI information to WDMAC electronically in an ASCII or Excel file.

The NDI file (File #9) will include cause of death information in the form of ICD-10 codes, and identifying information that will need to be stripped from the file before transmittal to WDMAC. Sites should have documentation describing the format of File #9, titled "Revised File Format for Cause of Death File #9," revised July 23, 1999, pages 11-17. If you do not have this documentation, please contact WDMAC. In order to prepare the electronic data file from NDI, sites should follow the below steps:

- Begin with the "File #9," one of the electronic data files received from NDI. If you have submitted multiple data sets to NDI and have received several File #9's, concatenate them and delete the records for the non-matches.
- Ensure that only true matches are included.
- Ensure that all participant identifying information is stripped from the file prior to sending it to WDMAC (see tables below).

DELETE:

Fields	Columns
1-13	1-100
16-17	120-126
19-30	133-164

KEEP:

Fields	Columns
14-15	101-119
18	127-132
31-83	165-438

Data collected through the NDI should be sent to WDMAC via e-mail as it is received. After receipt, WDMAC will review the NDI data to ensure that a *DENR* has been completed for each death discovered via NDI match. Sites will be notified of any missing *DENR* forms.

b. Local Death Registry

For matches obtained through local death registry searches, after completion of the *DENR* as described above, sites will need to complete and data enter a CORE Form to report the death information.

2. TB, CANCER AND AIDS REGISTRY MATCHES

The data received from TB, Cancer and AIDS registries will most likely be in a non-standard format depending on whether the match was done by computer or manually. When a match is made through the AIDS registry for one of the WIHS ascertainable events, only a *CORE* form need be completed to report the event. For TB and Cancer registry matches, two steps must be followed to report the event: (1) complete and data enter the appropriate registry confirmation form (*Tuberculosis – Verified Case Report* or *Cancer Registry Case Report*), and (2) complete and data enter a *CORE* form.

If knowledge of a reportable event is obtained through registry match, no further review of medical records need be done to verify the event. No *Ascertainment Tracking Checklist (ATC)* will be generated or completed to report TB, Cancer or AIDS registry matches. Since no *ATC* will be generated or completed, the event tracking number (ETN) will be left blank on *CORE* forms completed for TB, Cancer or AIDS registry matches.

When completing a *CORE* form for an event confirmed through registry match, the method of diagnosis should be coded as "no confirmation/clinician report." The level of confidence in the diagnosis will therefore be recorded as "indeterminate" since there was no actual review of medical records by WIHS CORE Team members.

NOTES:

- As noted in Communication Memo #384: All WIHS sites should prepare Cancer registry match submissions to include all women who have transferred from their site to another, as well as those currently at their site.
- A *Cancer Registry Case Report* should be completed for all in situ carcinomas reported through the Cancer registry. Ductal and lobular carcinomas in situ (breast) will also be reported on a *CORE* form. However, *CORE* forms will not be completed for reports of cervical carcinoma in situ (CIS), CIN2 or CIN3.
- Cancer events that occurred greater than five years before the baseline visit will only be reported through registry match – no medical record abstraction should be done for these cancers.
- General "AIDS" (disease code 220) and CD4 < 200 or CD4% < 14% (disease code 245) are collected and reported only when obtained from registry match.
- Disease code 170 (MAI/MAC) should be reported on the *CORE* form when a registry match is made for the combined events "MAI/MAC/M. Kansasii." Disease code 165 (CMV elsewhere in body) should be reported when a registry match is made for "Disseminated CMV." Disease code 124 (TB elsewhere in body) should be reported when a registry match is made for "extra-pulmonary TB." Disease code 175 (cryptococcal infection in blood or elsewhere) should be reported when a registry match is made for "cryptococcal infection."
- After each registry match, please send WDMAC an electronic file containing the list of WIHSID numbers that were searched. Also send to WDMAC the date of the match, the kind of the registry (TB, Cancer, etc.) that was searched, and whether the registry search was city, county, or state-wide.

- Beginning in 2008, cancer treatment data are to be obtained from the cancer registries. Since these data were not previously recorded on the WIHS cancer report form (*CNCR*), a new form must be completed for both previously reported and newly reported cases. Please review WIHS Communication Memo #510 for instructions on the completion and data entry of *CNCR* forms for participants with previously reported cancers.

For more information about reporting events found through registry matches on the *CORE* form, see the *WIHS Outcomes Ascertainment Protocol* (**Manual of Operations, Section 12**).

In addition, see the *WIHS Criteria for Clinical Diagnosis* (**Manual of Operations, Section 12, Appendix A**) and the *Disease Code List* (**Manual of Operations, Section 12, Appendix B**) for a complete description of the WIHS ascertainable events.

APPENDIX A: DRAFT PROTOCOL FOR COMPUTERIZED MATCH (modify as appropriate for your site)

Women's Interagency HIV Study (WIHS) & Alameda County TB Control Program

Background

The Alameda County Public Health Department maintains a registry of persons diagnosed with TB in Alameda. Persons with TB are reported to the health department by health care providers and by departmental staff who investigates suspected cases at hospitals and outpatient facilities. Detailed medical, sociodemographic, risk group, and medication information is ascertained through Chart review, in consultation with the medical provider. For TB cases, all diagnostic test results and sites of infection are recorded. Utilization and susceptibility of therapeutic medication is also maintained.

The Women's Interagency HIV Study is a multi-site, prospective study of HIV-infected women and women at high risk of HIV infection. The San Francisco Consortium is one of six national WIHS sites and is conducted by investigators at the University of California at San Francisco. Women were recruited from ongoing epidemiologic studies of HIV infection, from HIV medical care clinics, and through street outreach workers. The study sample in the San Francisco Bay Area consists of 90 HIV-uninfected and 340 HIV-infected women.

Women who agree to participate complete an interviewer-administered questionnaire regarding sociodemographic characteristics, risk behaviors, and medical history, followed by a complete physical examination. At the time of examination, specimens for detection of *Neisseria gonorrhoeae, Chlamydia trachomatis, Trichomonas vaginalis,* bacterial vaginosis, human papillomavirus, oral and vaginal candida, and a Papanicolaou smear are obtained. Serologic tests for HIV antibodies, HIV viral load, syphilis, hepatitis A, B, C, complete blood cell count, t lymphocyte subsets, and pregnancy are also obtained. Annual PPD and anergy skin tests are performed. Data is collected at baseline and at six-month intervals. Medical chart reviews are conducted for any participants who acknowledge an AIDS-indicator condition, cancer, TB, or previous hospitalization.

The establishment of a match between the Alameda TB Control Program registry and participants in WIHS would provide an additional source of TB case reporting for Alameda County. The detailed medical information contained in the TB case registry would augment the information collected from WIHS participant interviews and medical chart reviews. In Addition, comparison of the medical history data in WIHS to the TB case registry would allow for validation of self-reported information.

Objectives

1. To identify women meeting the current surveillance definition of TB who have not been reported to Alameda TB Control.
2. To augment medical information and vital status for WIHS participants.
3. To measure the validity of self-reported medical information.

Matching Definition

Matching variables are defined as primary and secondary. Primary matching variables are name, date of birth, and social security number. Name is considered to be a perfect match if: a) first and last names match exactly, or b) last name matches exactly and first names are functional equivalents (e.g. Susan and Sue). Secondary matching variables are race and address.

Matches will be classified as Definite Matches and Non-matches. Matches will be considered definite if all three primary matching variables match perfectly or if two of the three primary variables match perfectly and the non-perfect matching variable is close (e.g. a transposition of numbers or letters) or is missing from one of the data files; and one of the two secondary matching variables match perfectly. Non-matches include those situations where none of the three primary variables match. Matches that do not meet the criteria for definite or non-matches will have hard copies and, if necessary, medical charts

reviewed to identify additional information or evaluate conflicting data to determine if the case is a definite match. Cases that do not reveal strong evidence for a match will be considered non-matches.

Matching Hardware, Software, and Analyses

All matching and analyses will be performed on a personal computer at the TB control office. Identifying information from the TB registry is maintained in an encrypted file. This file will be merged with the non-identifying information from the TB registry and converted into a SAS file. These data will be stored on a computer diskette that will be removed at the end of each work and stored in a locked safe. A SAS data file of WIHS information will be provided to the TB Control Office. This file will contain both the matching variables and other data of interest (e.g. information necessary to determine risk group and TB diagnosis and medications. Files from WIHS and the TB case registry will be merged into one data file for analyses.