

# WIHS Public Data Set

## A. General Information

Release Information: WDMAC (The WIHS Data Management and Analysis Center) released the first WIHS public data set in March 2002. Future data will be released on a yearly schedule with new releases superceding previous versions. The current release (P04) covers data collected from visit 1 (October 1, 1994) through visit 15 (March 31, 2002).

Data Summary: 1) The cutoff date for data contained in each release is March 31 three years preceding the current WIHS visit. New variables, reflecting future research initiative additions, will be incorporated in each new release of the public data set. 2) Data components cover all visit data collected. 3) Information that might be used to identify individuals has been modified; for example, the WIHS participant identification number (WIHSID) is replaced by a randomized identification number, information pertaining to study site is deleted, all dates are stripped of day and month value, retaining only the year value.

Data Manipulation: The original WIHSID number contains information that could be used to link a participant to the center and subsite at which she was recruited. Therefore, WIHSID has been replaced by a randomized identification number for each of the 2,625 participants. Once the data sets were created, the original WIHSIDs were deleted and the randomized IDs were sorted in ascending order. Additionally, the day and month of each participant's date of birth were deleted, leaving only the year of birth, and the dates of visit, event report, clinical diagnosis, hospitalization and death retain only year values, as well.

Before data were released to the public, 10 percent of the variables from each public data file were sampled and compared with the same set of variables in the raw data files for each visit. All data files and their associated codebooks and SAS input statement files were also copied to the distribution media. The "README" file distributed with the set includes information describing the data files and the cutoff date of the event files. All questionnaire forms are included to facilitate understanding of the codebook for each data file.

Data Medium: The current release of the public data set was submitted on a CD-ROM disc in ASCII text of PC/Windows structure.

## B. Ordering Information

To ORDER the WIHS Public Data Set, Release P04, contact Ms. Cathy Metzler of NTIS at (703) 605-6129 or Ms. Brenda Burba at (703) 605-6004; or go to the [NTIS](#) web page. Please reference the NTIS order number PB-2005-500136. The cost of the CD is \$79.00 (plus \$5.00 for shipping and handling).

## C. Associated Questionnaire Forms and Documentation

Forms: Five main types of data collection forms were submitted with the data set. 1) Central laboratory results are reported on forms beginning with the letter “C,” e.g., C45 – Bacterial Vaginosis Smear Gram Stain. 2) Local laboratory results are reported on forms beginning with the letter “L,” e.g., L01 – HIV ELISA and Western Blot. 3) Interview and physical exam forms, both baseline and follow-up, begin with the letter “F,” e.g., F06 – Baseline Psychosocial Measures and F26 – Follow-up Psychosocial Measures. 4) Oral Substudy forms begin with “OP,” e.g., OP04 – Oral Mucosal Tissue Exam. 5) Medication usage forms begin with “DRUG,” e.g., DRUG 1 – Antiviral Medications.

Additionally, there are five miscellaneous files included in the data set. 1) ABRV – Abbreviated Visit, for those women incarcerated or too sick to complete a full visit. 2) INT – Interim Events Reporting Form. 3) MENT – Mental Alternations Test. 4) HIVSTAT – a summary file containing sero-status information on all participants. 5) Outcome – a summary file containing clinical outcome data.

The majority of the forms are in Microsoft Word format, with the exception of drug1.wpd, drug2.wpd, f29a.wpd, int.wpd and l15.wpd, which are in WordPerfect.

Other Documentation: Archives of WIHS publications and a directory of WIHS investigators can be accessed from the WIHS home page (<http://statepiaps.jhsph.edu/wihs/>).

## D. List of Files in the Release

### Name – Description

- README – The README file accompanying the data
- ABRV – Abbreviated Visit
- ABRVs1 – Abbreviated Visit, subform 1: Medication Use
- ABRVs2 – Abbreviated Visit, subform 2: Problematic Questions
- C30 – Toxoplasmosis
- C31 – HTLV 1 and 2
- C45 – Bacterial Vaginosis Smear Gram Stain
- C50 – Urine for Chlamydia
- C51 – Chlamydia Swab Confirmatory
- C54 – HIV Viral Load
- C60 – Pap Smear
- C65 – Syphilis DFA – Genital Ulcers and Fissures
- C66 – Serum Herpes Serologies
- C70 – Stimulated Saliva Evaluation (Oral Substudy)
- C71 – Subgingival Plaque (Oral Substudy)
- C72 – Erythematous Candidiasis Results
- DRUG1 – Antiviral Medication Use
- DRUG2 – Non-antiviral Medication Use
- F00 – Screening Form

- F01 – Baseline Sociodemographics
- F02 – Baseline Medical and Health History
- F02s1 – Baseline Medical and Health History, subform 1: Research Study Medication Use
- F02s2 – Baseline Medical and Health History, subform 2: Other Medication Use
- F03 – Baseline Obstetric, Gynecologic and Contraceptive History
- F03s1 – Baseline Obstetric, Gynecologic and Contraceptive History, subform 1: Pregnancy History
- F04 – Baseline Alcohol, Drugs and Sexual Behavior
- F05 – Baseline Health Care Utilization
- F06 – Baseline Psychosocial Measures
- F06s1 – Baseline Psychosocial Measures, subform 1: Deceased Children
- F06s2 – Baseline Psychosocial Measures, subform 2: Living Children
- F06s3 – Baseline Psychosocial Measures, subform 3: Interviewer Assessment
- F07 – Baseline and Follow-up Physical Exam
- F07s1 – Physical Exam, subform 1: Skin Lesions
- F07s2 – Physical Exam, subform 2: Oral Lesions
- F08 – Baseline and Follow-up Gynecological Exam
- F08s1 – Gynecological Exam, subform 1: Abnormal Lesions
- F08A – Potential CVL Contaminants
- F09 – Baseline Blood Specimen Collection
- F10 – Plasma and Cell Separation and Freezing
- F11 – Baseline Specimens Collection During Physical Exam
- F21 – Follow-up Sociodemographics
- F22 – Follow-up Medical and Health History, part 1 (start – question E8)
- F22p1 – Follow-up Medical and Health History, part 2 (questions E9 – E23)
- F22p2 – Follow-up Medical and Health History, part 3 (question F1 – end)
- F22s1 – Follow-up Medical and Health History, subform 1: Research Study Medication Use
- F22s2 – Follow-up Medical and Health History, subform 2: Other Medication Use
- F22p2s1 – Follow-up Medical and Health History, part 3, subform 1: Antiretroviral Medication Use
- F22p2s2 – Follow-up Medical and Health History, part 3, subform 2: Non-antiretroviral Medication Use
- F22p2s3 – Follow-up Medical and Health History, part 3, subform 3: Alternative and Complementary Medication Use
- F22p2s4 – Follow-up Medical and Health History, part 3, subform 4: Other Prescription Medication Use
- F22p2s5 – Follow-up Medical and Health History, part 3, subform 5: Other Symptoms Report
- F23 – Follow-up Obstetric, Gynecologic and Contraceptive History
- F23s1 – Follow-up Obstetric, Gynecologic and Contraceptive History, subform 1: Pregnancy History
- F24 – Follow-up Alcohol, Drug and Sexual Behavior
- F25 – Follow-up Health Care Utilization

- F25A – Follow-up Health Care Utilization Supplement
- F26 – Follow-up Psychosocial Measures
- F26s1 – Follow-up Psychosocial Measures, subform 1: Deceased Children
- F27 – Karnofsky Scale
- F27s1 – Karnofsky Scale, subform 1: Problematic Questions
- F29 – Follow-up Blood Specimen Collection Form
- F29a – Antiretroviral Usage Assessment for Blood Draw
- F29as1 – Antiretroviral Usage Assessment for Blood Draw, subform 1: Antiretroviral Medication Use in Past 3 Days
- F29as2 – Antiretroviral Usage Assessment for Blood Draw, subform 2: Date/Time of Antiretroviral Medication Use
- F31 – Follow-up Specimens Collected During Physical Exam
- HIVSTAT – Sero-status summary file
- INT – Interim Events Reporting Form
- L01 – HIV ELISA and Western Blot
- L02 – Serum Hepatitis
- L03 – Automated CBC/Differential
- L03A – Hand/Manual Differential
- L04 – Flow Cytometry
- L05 – Liver/Renal Function Tests
- L06 – Serum Syphilis
- L07 – Anergy Panel
  
- L08 – PPD Skin Test
- L09 – Chlamydia Gen-Probe
- L10 – Urinalysis
- L11 – Urine Culture Result
- L12 – (Urine) Pregnancy Test
- L13 – Gonorrhea Gen-Probe
- L14 – Colposcopy Results
- L15 – Pelvic Exam Histopathology Report
- L16 – Dysplasia Treatment
- L17 – HSV Culture of Ulcers and Fissures
- L18 – Trichomonas Vaginalis Culture
- L19 – CVL Processing
- MENT – Mental Alternations Test
- OP01 – Medical Evaluation
- OP02 – Research Interview
- OP04 – Saliva Sample Collection
- OP04 – Oral Mucosal Tissue Exam
- OP04s1 – Oral Mucosal Tissue Exam, subform 1: Lesions
- OP05 – Smear Results
- OP06 – Tooth Count and Random Half Mouth
- OP07 – Plaque Index
- OP08 – Gingival Banding Score

- OP09 – Papillary Assessment
- OP10 – Subgingival Plaque
- OP11 – Coronal Caries
- OP12 – Root Caries
- OP13 – Gingival Bleeding
- OP14 – Loss of Attachment, part 1
- OP14p1 – Loss of Attachment, part 2
- OP15 – Dental Prosthesis
- Outcome – Clinical Outcome summary file