**GENOMICS DATA USE CERTIFICATION AGREEMENT**

**Women’s Interagency HIV Study**

***(January 17, 2017 version)***

**INTRODUCTION AND STATEMENT OF POLICY**

The Women’s Interagency HIV Study (WIHS), funded by the National Institutes of Health (NIH), has an established central host genomic data repository called the WIHS Genomics Database for securely storing and sharing human host genomic data submitted to the WIHS. Implicit in the establishment of the WIHS Genomics Database is that scientific progress in genomic research will be greatly enhanced if the data are readily available to all approved scientific investigators and shared in a manner consistent with the research participants’ informed consent.

Access to human genomic data will be provided to research investigators who, along with their WIHS Contact (e.g., site PI), have certified their agreement with the expectations and terms of access detailed below. It is the intent of NIH and the WIHS that approved users of the WIHS Genomics Database datasets recognize any restrictions on data use established by the WIHS stated on the WIHS Concept Sheet, the WIHS Genomic Data Code of Conduct, and the WIHS Genomic Data Use Contract. Definitions of terminology used in this document are found in Appendix 1.

This **Data Use Agreement** (the “Agreement”) is by and between the WIHS Principal Investigators [Drs. Igho Ofotokun and Gina Wingood (Emory University); and Dr. Mary Young (Georgetown University); and Drs. Stephen Gange and Elizabeth Golub (Johns Hopkins University); and Dr. Kathryn Anastos (Montefiore Medical Center); and Drs. Mardge Cohen and Audrey French (Chicago WIHS Consortium : Cook County Hospital, University of Illinois at Chicago, Rush-Presbyterian-St. Luke's Medical Center, and Northwestern Memorial Hospital); and Drs. Michael Saag and Mirjam Colette Kempf (The University of Alabama at Birmingham); and Drs. Ruth Greenblatt, Phyllis Tien and Bradley Aouizerat (The University of California at San Francisco); and Drs. Margaret Fischl and Lisa Metsch (The University of Miami); and Dr. Deborah Konkle- Parker (The University of Mississippi at Jackson); and Dr. Ada Adimora (The University of North Carolina at Chapel Hill); and Drs. Howard Minkoff and Deborah Gustafson (The State University of New York at Brooklyn), (hereinafter referred to individually as “WIHS CONSORTIUM”)and [**INSERT INVESTIGATOR NAME]** (“Approved User”). This agreement (“Agreement”) establishes the terms and conditions under which the User will obtain and use certain data (“Genomics Dataset”) from the Women’s Interagency HIV Study (WIHS) maintained by the WIHS Consortium.

The effective date of this agreement shall be the Concept Sheet Approval Date, as specified on the WIHS Forum and WIHS Data Management and Analytic Center (WDMAC) approval notification.

**TERMS OF ACCESS**

**1. Research Use**

The Requester (e.g., Lead Investigator) agrees that if access is approved, (1) the PI named in the WIHS Concept Sheet and/or (2) those named in the “Lead Investigator” section A.5. of the WIHS Concept Sheet, and any trainee, employee, or contractor[[1]](#footnote-1) working on the proposed research project under the direct oversight of these individuals, shall become Approved Users of the requested dataset(s). Research use will occur solely in connection with the approved research project described in the WIHS Concept Sheet, which includes a description of the research objectives and design. New uses of these data outside those described in the WIHS Concept Sheet will require submission of a new WIHS Concept Sheet; modifications to the research project will require submission of an amendment to this application (e.g., adding or deleting collaborators from the same institution, adding datasets to an approved project). Access to the requested dataset(s) is granted for a period of 1 year as defined below.

**2. Requester and Approved User Responsibilities**

The Requester (e.g., Lead Investigator) agrees through the submission of the WIHS Concept Sheet that the Investigators named in the WIHS Concept Sheet have reviewed and understand the principles for responsible research use and data handling of the genomic datasets as defined in WIHS Genomics Data Use Contract, the WIHS Genomics Data Code of Conduct, and the WIHS Data Security Best Practices Requirements. The Requester (e.g., Lead Investigator) and Approved Users (e.g., co-Investigators, trainees, Contractors) further acknowledge that they are responsible for ensuring that all uses of the data are consistent with federal, state, and local laws and regulations and any relevant institutional policies. The Requester certifies that all investigators are in good standing (i.e., no known sanctions) with the institution, relevant funding agencies, and regulatory agencies and is eligible to conduct independent research (i.e., is not a postdoctoral fellow, student, or trainee). The Requester and all Approved Users may use the dataset(s) only in accordance with the parameters described in the approved WIHS Concept Sheet for the appropriate research use, as well as any limitations on such the use, of the dataset(s) and as described in the approved WIHS Concept Sheet and as required by law.

Through submission of the WIHS Concept Sheet, the Lead Investigator agrees to submit either a project renewal or close-out request prior to the expiration date of the 1-year data access period. The Lead Investigator also agrees to submit an annual progress update or a final progress report at the 1-year anniversary of the approved WIHS Concept Sheet, as described under *Research Use Reporting* below. Failure to submit a renewal or complete the close-out process, including confirmation of data destruction by the WIHS PI/Contact, may result in termination of all current data access and/or suspension of the Lead Investigator and all associated key personnel and collaborators from submitting new WIHS Concept Sheets for a period to be determined by NIH. Repeated violations or unresponsiveness to WIHS and/or NIH requests may result in further measures affecting the Requester.

Approved Users who may have access to personal identifying information for research participants in the original study at their institution or through their collaborators may be required to have IRB approval. By approving and submitting the attached WIHS Concept Sheet, the WIHS PI/Contact provides assurance that relevant institutional policies and applicable federal, state, and local laws and regulations (if any) have been followed, including IRB approval if required. The WIHS PI/Contact also assures through the approval of the WIHS Concept Sheet that other institutional departments with relevant authorities (e.g., those overseeing human subjects research, information technology, or technology transfer) have reviewed the relevant sections of the WIHS Genomics Data Use Contract, the WIHS Genomics Data Code of Conduct, and the WIHS Data Security Best Practices Requirements and the associated procedures and are in agreement with the principles defined.

In some cases, NIH anticipates that the WIHS Genomics Database datasets will be updated with additional information. Unless otherwise indicated, all statements herein are presumed to be true and applicable to the access and use of all versions of these datasets.

**3. Non-Identification**

Approved Users agree not to use the requested datasets, either alone or in concert with any other information, to identify or contact individual participants from whom data and/or DNA samples were collected. This provision does not apply to research investigators operating with specific IRB approval, pursuant to 45 CFR

46, to contact individuals within datasets or to obtain and use identifying information under an IRB approved research protocol. All investigators conducting “human subjects research” within the scope of 45 CFR 46 must comply with the requirements contained therein.

**4. Non-Transferability**

The Requester and Approved Users agree to retain control of the data and further agree not to distribute data obtained through the WIHS Concept Sheet to any entity or individual not covered in the approved WIHS Concept Sheets. If Approved Users are provided access to WIHS genomic datasets for inter-institutional collaborative research described in Section B of the WIHS Concept Sheets, and all members of the collaboration are also Approved Users through their home institution(s), data obtained through this WIHS Concept Sheet may be securely transmitted within the collaborative group. Approved Users are expected to follow all data security practices and other terms of use defined in this agreement and the dbGaP Security Best Practices (see Appendix 2 and 3) for raw data and any derived data, including transmission of these data.

The Requester and Approved Users acknowledge responsibility for ensuring the review and agreement to the terms within this WIHS Concept Sheet and the appropriate research use of WIHS genomic data by research staff associated with any approved project, subject to applicable laws and regulations. WIHS genomic datasets obtained through this WIHS Concept Sheet, in whole or in part, may not be sold to any individual at any point in time for any purpose.

Investigators agree that if they change institutions during the access period they will complete the WIHS Concept Sheets close-out process before moving to their new institution or submit a modification of the approved WIHS Concept Sheet. A new WIHS Concept Sheet, in which the new institution agrees to the NIH WIHS Data Sharing Policy, may be required before data access resumes. The requirement of a new WIHS Concept Sheet will be determined by the WIHS EC. As part of the close-out process, any versions of the data stored at the prior institution should be destroyed and destruction confirmed in writing by the Concept Sheet PI and the WIHS PI/Contact, as described below. However, with advance written notice and approval by the WIHS EC to transfer responsibility for the approved research project to another Approved User from the PI’s prior institution, the data may not need to be destroyed.

**5. Data Security and Data Release Reporting**

The Requester and Approved Users acknowledge the NIH’s/WIHS’s expectation that they have reviewed and agree to handle the requested dataset(s) according to the current dbGaP Security Best Practices (see Appendix 2 and 3), including its detailed description of requirements for security and encryption. These include, but are not limited to:

* All Approved Users have completed all required computer security training required by their institution, for example, the <http://irtsectraining.nih.gov/>, or the equivalent.
* The data will always be physically secured (e.g., through camera surveillance, locks on doors/computers, security guard).
* Servers must not be accessible directly from the internet, (e.g., they must be behind a firewall or not connected to a larger network) and unnecessary services should be disabled.
* Use of portable media (e.g., CD, flash drive or laptop) is discouraged, but if necessary then they must be encrypted consistent with applicable law.
* Updated anti-virus/anti-spyware software is used.
* Security auditing/intrusion detection software that regularly scans and detects potential data intrusions must be in place.
* Strong password policies for file access are used.
* All copies of the dataset are destroyed, as permitted by law and local institutional policies, whenever any of the following occurs:
  + the WIHS Concept Sheet expires and renewal is not sought;
  + access renewal is not granted;
  + WIHS EC requests destruction of the dataset; and
  + Continued use of the data would no longer be consistent with the WIHS Concept Sheet.

In addition, the Requester and Approved Users agree to keep the data secure and confidential at all times and to adhere to information technology practices in all aspects of data management to assure that only authorized individuals can gain access to WIHS genomic datasets. This agreement includes the maintenance of appropriate controls over any copies or derivatives of the data obtained through this Data Access Request.

Requesters and PIs agree to notify the WIHS of any unauthorized data sharing, breaches of data security, violations in the presentation and publication embargo period, or inadvertent data releases that may compromise data confidentiality within 24 hours of when the incident is identified. As permitted by law, notifications should include any known information regarding the incident and a general description of the activities or process in place to define and remediate the situation fully. Within 3 business days of the WIHS EC, the Requester, through the PI and the WIHS PI/Contact, agree to submit to the WDMAC a more detailed written report including the date and nature of the event, actions taken or to be taken to remediate the issue(s), and plans or processes developed to prevent further problems, including specific information on timelines anticipated for action.

**All notifications and written reports of data security incidents should be sent to:**

WIHS Data Management and Coordinating Center (WDMAC) ([jhsph.wdmac@jhu.edu](mailto:jhsph.wdmac@jhu.edu)).

WIHS, NIH, or another entity designated by NIH may, as permitted by law, also investigate any data security incident. Approved Users and their associates agree to support such investigations and provide information, within the limits of applicable local, state, and federal laws and regulations. In addition, Requesters and Approved Users agree to work with the WIHS and NIH to assure that plans and procedures that are developed to address identified problems are mutually acceptable and consistent with applicable law.

**6. Intellectual Property**

By requesting access to genomic dataset(s), the Requester and Approved Users acknowledge the intent of the NIH and WIHS that anyone authorized for research access through the attached approved WIHS Concept Sheet follow the intellectual property (IP) principles in the [NIH GWAS Data Sharing Policy](http://grants.nih.gov/grants/guide/notice-files/NOT-OD-07-088.html#property) as summarized below:

Achieving maximum public benefit is the ultimate goal of data distribution through the WIHS Genomic Database repository. The WIHS encourages a responsible approach to management of intellectual property derived from downstream discoveries, as outlined in the NIH’s [Best Practices for the Licensing of Genomic Inventions](http://www.ott.nih.gov/policy/genomic_invention.html) and its [Research Tools Policy.](http://ott.od.nih.gov/policy/research_tool.html)

The NIH considers these data as pre-competitive and urges Approved Users to avoid making IP claims derived directly from the genomic dataset(s). It is expected that these NIH-provided data, and conclusions derived therefrom, will remain freely available, without requirement for licensing. However, the NIH also recognizes the importance of the subsequent development of IP on downstream discoveries, especially in therapeutics, which will be necessary to support full investment in products to benefit the public.

**7. Research Dissemination and Acknowledgement of WIHS Genomic Study Datasets**

It is WIHS’s intent to promote the dissemination of research findings from WIHS genomic dataset(s) as widely as possible through scientific publication or other appropriate public dissemination mechanisms. Approved Users are strongly encouraged to publish their results in peer-reviewed journals and to present research findings at scientific meetings.

In accord with the [WIHS Publication Policy](https://statepi.jhsph.edu/wihs/admin/moo/), and as expressed through the submission of the WIHS Concept Sheet and Data Use Agreement, Approved Users acknowledge NIH’s and WIHS’s expectation that they will not submit findings using the WIHS dataset(s), or updated versions thereof, for publication or presentation without approval of the WIHS (i.e., following review of said publication or presentation by WIHS co-authors).

Approved Users agree to acknowledge the WIHS Contributing Investigator(s) who submitted data from the original study, the primary funding organization that supported the Contributing Investigators, and the WIHS data repository, in all oral and written presentations, disclosures, and publications resulting from any analyses of WIHS data. A sample short acknowledgment statement for the dataset(s) follows, and a detailed explanation of these procedures can be found in the [WIHS Publication Policy](https://statepi.jhsph.edu/wihs/admin/moo/).

*Data in this manuscript were collected by the Women’s Interagency HIV Study (WIHS). The contents of this publication are solely the responsibility of the authors and do not represent the official views of the National Institutes of Health (NIH). WIHS (Principal Investigators): UAB-MS WIHS (Michael Saag, Mirjam-Colette Kempf, and Deborah Konkle-Parker), U01-AI-103401; Atlanta WIHS (Ighovwerha Ofotokun and Gina Wingood), U01-AI-103408; Bronx WIHS (Kathryn Anastos), U01-AI-035004; Brooklyn WIHS (Howard Minkoff and Deborah Gustafson), U01-AI-031834; Chicago WIHS (Mardge Cohen and Audrey French), U01-AI-034993; Metropolitan Washington WIHS (Seble Kassaye), U01-AI-034994; Miami WIHS (Margaret Fischl and Lisa Metsch), U01-AI-103397; UNC WIHS (Adaora Adimora), U01-AI-103390; Connie Wofsy Women’s HIV Study, Northern California (Ruth Greenblatt, Bradley Aouizerat, and Phyllis Tien), U01-AI-034989; WIHS Data Management and Analysis Center (Stephen Gange and Elizabeth Golub), U01-AI-042590; Southern California WIHS (Joel Milam), U01-HD-032632 (WIHS I – WIHS IV). The WIHS is funded primarily by the National Institute of Allergy and Infectious Diseases (NIAID), with additional co-funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), the National Cancer Institute (NCI), the National Institute on Drug Abuse (NIDA), and the National Institute on Mental Health (NIMH). Targeted supplemental funding for specific projects is also provided by the National Institute of Dental and Craniofacial Research (NIDCR), the National Institute on Alcohol Abuse and Alcoholism (NIAAA), the National Institute on Deafness and other Communication Disorders (NIDCD), and the NIH Office of Research on Women’s Health. WIHS data collection is also supported by UL1-TR000004 (UCSF CTSA) and UL1-TR000454 (Atlanta CTSA).*

**8. Research Use Reporting**

To assure adherence to NIH and WIHS policies and procedures for genomic data, PIs agree to provide annual progress updates on how these data have been used, including presentations, publications, and the generation of intellectual property. This information helps WIHS evaluate program activities and may be considered by the WIHS governance committees as part of NIH’s effort to provide ongoing oversight and management of WIHS genomic data sharing activities.

As stipulated in the WIHS Data Use Agreement, progress updates must be provided as part of the annual project renewal or project close-out processes, prior to the expiration of the 1-year data access period. PIs who are seeking renewal or close-out of a project agree to complete the appropriate online forms and provide specific information such as publications or presentations that resulted from the use of the requested dataset(s), a summary of any plans for future research use, any violations of the terms of access described within this DUA and the implemented remediation, and information on any downstream intellectual property generated from the data. PIs also may include general comments regarding topics such as the effectiveness of the data access process (e.g., ease of access and use), appropriateness of data format, challenges in following the policies, and suggestions for improving data access or the program in general.

*Note that any inadvertent or inappropriate data release incidents should be reported to WDMAC (*[*jhsph.wdmac@jhu.edu*](mailto:jhsph.wdmac@jhu.edu)*) according to the agreements and instructions under Term 5.*

**9. Non-Endorsement, Indemnification**

The Requester and Approved Users acknowledge that although all reasonable efforts have been taken to ensure the accuracy and reliability of WIHS genomic data, the NIH, the WIHS, and WIHS Contributing Investigators do not and cannot warrant the results that may be obtained by using any data included therein. NIH, the WIHS, and all contributors to these datasets disclaim all warranties as to performance or fitness of the data for any particular purpose.

No indemnification for any loss, claim, damage, or liability is intended or provided by any party under this agreement. Each party shall be liable for any loss, claim, damage, or liability that said party incurs as a result of its activities under this agreement, except that NIH, as an agency of the United States, may be liable only to the extent provided under the Federal Tort Claims Act, 28 USC 2671 et seq.

**10. Termination and Violations**

This DUA will be in effect for a period of 1 year from the date the dataset(s) are made accessible to the PI (“Approved Access Date”). At the end of the access period, PIs agree to report progress, and renew access or close-out the project. Upon project closure-out, all Approved Users agree to destroy all copies of the requested dataset(s), except as required by publication practices or law to retain them. Copies of NIH genomic dataset(s) may not need to be destroyed if, with advance notice and approval by the WDMAC, the project has been transferred to another Approved User at the same institution. In this case, documentation must be provided that other Approved Users are using the dataset(s) under an approved WIHS Concept Sheet.

**The Requester and PI acknowledge that the NIH or the WIHS may terminate this agreement and immediately revoke access to all WIHS genomic datasets at any time if the Requester is found to be no longer in agreement with the policies, principles and procedures of the NIH and the WIHS**.

**BY SUBMISSION OF THE GENOMIC DATA USE AGREEMENT (DUA):**

* The Requester through the WIHS PI/Contact attests to the Concept Sheet PI’s qualifications for access to and use of WIHS genomic dataset(s) and agrees to the NIH and WIHS principles, policies, and procedures for the use of the requested datasets as articulated in this document and as summarized in the WIHS Concept Sheet and Data Use Agreement, including the potential termination of access should any of these terms be violated.
* Requesters and the PI further acknowledge that they have shared this document, the WIHS Approved User Code of Conduct, and the WIHS Genomics Data Sharing Policy, and procedures for access and use of genomic datasets with any Approved Users, appropriate research staff, and all other Key Personnel and collaborators identified in the WIHS Concept Sheet.
* WIHS PI/Contact acknowledges that they have considered the relevant NIH and WIHS data sharing policies and procedures, that they have shared this document and the relevant policies and procedures with appropriate institutional departments, and have assured compliance with local institutional policies related to technology transfer, information technology, privacy, and human subjects research.
* WIHS PI/Contact also acknowledges that their institute is solely responsible for the conduct of all individuals who have access to the data under the WIHS Concept Sheet, including investigators, contractor staff (both on and off-site) and trainees, as specified in the WIHS Data Use Agreement (DUA).

|  |  |
| --- | --- |
| WIHS Readme |  |
| Project title |  | |
| Printed Name and Title: | [Lead Investigator] | [WIHS Site PI or WIHS Affiliate] |
| Signature |  |  |
| Date |  |  |
| Contact Information: |  |  |
| Address: |  |  |
| Phone: |  |  |
| Fax: |  |  |
| Email: |  |  |

**APPENDIX 1: DEFINITIONS**

**Approved User:** A user approved by the WIHS to access one or more datasets for a specified period of time and only for the purposes outlined in the PI’s approved research use statement. Staff members and trainees under the direct supervision of the PI are also Approved Users and must abide by the terms laid out in the Data Use Contract agreement.

**Collaborator:** An individual who is not under the direct supervision of the principal investigator (PI) (e.g., not a member of the PI’s laboratory) who assists with the PI’s WIHS Genomic research project. Internal collaborators are employees of the Requester and work at the same location/campus as the PI. External collaborators are not employees of the Requester and/or do not work at the same location as the PI, and consequently must be independently approved to access WIHS data.

**Contributing Investigator:** An investigator who submitted a genomic dataset to the WIHS data repository (e.g., WIHS Genomic Database).

**Data Access Request (DAR):** A request submitted to WIHS for a specific “consent group” specifying the data to which access is sought, the planned research use, and the names of collaborators and the Information Technology Director. The DAR is signed by the investigator requesting the data and her/his WIHS PI/Contact. Collaborators and project team members on a request must be from the same institution or organization.

**Data Derivative:** any data including individual-level data or aggregate genomic data that stems from the original dataset deposited (e.g. imputed or annotated data) in NIH-designated data repositories (e.g., WIHS Genomic Database). Summary information that is expected to be shared through community publication practices in not included in this term.

**Data Use Agreement (DUA):** An agreement between the Approved Users, the Requestor, and WIHS regarding the terms associated with WIHS Genomic Database data access and the expectations for use of WIHS datasets.

**WIHS Approved User Code of Conduct:** Key principles and practices agreed to by all research investigators requesting access to NIH controlled-access genomic data. The elements within the Code of Conduct reflect the terms of access in the Data Use Certification agreement. Failure to abide by the Code of Conduct may result in revocation of an investigator’s access to any and all approved datasets.

**WIHS PI/Contact:** Generally, WIHS Site PI and/or recognized WIHS Investigator.

**Progress Update:** Information included with the annual WIHS Concept Sheet renewal or close-out summarizing the analysis of WIHS datasets obtained through the DAR and any publications and presentations derived from the work.

**Project Closeout:** Closeout of a research project that used controlled-access data from an NIH-designated data repository (e.g., WIHS Genomics Database) and confirmation of data destruction when the research is completed and/or discontinued.

**Project Renewal:** Renewal of a principal investigator’s access to controlled-access datasets for a prior-approved project before the expiration date of a WIHS Concept Sheet.

**APPENDIX 2. dbGaP SECURITY BEST PRACTICES**

# SECURITY BEST PRACTICES

# Introduction

The data sets provided in conjunction with this agreement are controlled access data. The procedures described below are based on the assumption that access to deidentified individual (person) level detailed genomic data associated with phenome data should be controlled and not publicly available.

The goal of this process is to ensure that data provided by the NIH is kept sufficiently secure and not released to any person not permitted to access the data, either through malicious or inadvertent means. To accommodate these requirements, systems housing these data must not be directly accessible from the internet, and the data must not be posted on any web or ftp server. Data placed on shared systems must be secured and limited to those involved in the research for which the data has been requested. If data is stored on laptops, mobile devices or removable media, those devices must be encrypted and password protected.

# Protecting the Security of Controlled Data

*Security Awareness Requirements*

This document provides an overview of security principles for data, access, and physical security to ensure confidentiality, privacy and accessibility of data.

The controlled access data you received is considered sensitive information. By following the **best practices** below, you will be doing much towards protecting the information entrusted to your care. This is a minimum set of requirements; additional restrictions may be needed by your institution and should be guided by the knowledge of the user community at your institution.

# Think Electronic Security

***The Single Most Important Advice:*** Download data to a secure computer or server and not to unsecured network drives or servers.

1. Make sure these files are never exposed to the Internet. Data must never be posted on a PI’s (or institution’s) website because the files can be “discovered” by internet search engines, e.g., *Google, Bing*.
2. Have a strong password for file access and never share it. A strong password has to meet at minimum following requirements:
   * At least 8 characters long
   * Does not contain your user name, real name, or company’s name
3. Does not contain complete dictionary word.
4. Contains characters from each of the following four groups: lowercase letters, uppercase letters, numerals, and special characters.
5. If you leave your office, close out your data files Or lock your computer.
   * Install a password-enabled screen saver that activates after 15 minutes of inactivity.
6. Data stored on laptops, mobile devices (e.g. smartphones, tablets, mp3 players) or removable media (e.g. CDs, flash drives, USB drives) must be encrypted. Most operating systems have the ability to natively run an encrypted file system or encrypt portions of the file system. (Windows = EFS or Pointsec and Mac OSX = File Vault)
7. Ensure that the system files and the registry are protected using strong access control lists (ACL), which are the lists of permissions attached to an object. ACLs or access permissions specify which user or process has an access to objects and what kind of operations are allowed on objects.
   * Appropriate use of ACLs will ensure limited access to downloaded data
8. Keep all software patches up-to-date

# Think Physical Security

1. If the data are in hard copy or reside on portable media, e.g., on a CD, flash drive or laptop), treat it as though it were cash.
2. Don’t leave it unattended or in an unlocked room.
3. Consider locking it up.
4. To maintain a secure environment, you must restrict physical access to all servers, network hardware, and backup media.
5. Exercise caution when traveling with portable media, i.e., take extra precautions to avoid the possibility of loss or theft (especially flash drives which are small and can easily be misplaced).

# Only Use Data by Approved Users on Secure Systems Protect the Security of Controlled Data on Servers

# Servers must not be accessible directly from the internet, (i.e. must be behind a firewall or not connected to a larger network) and unnecessary services disabled.

# Enforce principle of Least Privilege to ensure that individuals and/or processes granted only the rights and permissions to perform their assigned tasks and functions, but no more.

# Keep systems up to date with security patches.

# dbGaP data on the systems must be secured from other users (restrict directory permissions to only the owner and group) and if exported via file sharing, ensure limited access to remote systems.

# If accessing systems remotely, encrypted data access must be used (such as SSH or VPN). It is preferred to use a tool such as RDP, X-windows or VNC that does not permit copying of data and provides “View only” support.

# Ensure that all users of this data have IT security training suitable for this data access and understand the restrictions and responsibilities involved in access to this data.

# If data is used on multiple systems (such as a compute cluster), ensure that data access policies are retained throughout the processing of the data on all the other systems. If data is cached on local systems, directory protection must be kept, and data must be removed when processing is complete.

# Requesting Investigators must meet the spirit and intent of these protection requirements to ensure a secure environment 24 hours a day for the period of the agreement.

# Maintain source data, and control copies of data

# The requesting investigator must retain the original version of the encrypted data.

# The requesting investigator must track any copies or extracts made of the data and shall make no copy or extract of the subject data available to anyone except an authorized staff member for the purpose of the research for which the subject data were made available.

# Collaborating investigators from other institutions must complete an independent data use certification to gain access to the data.

# When use of the dataset is complete — destroy all individual level data

# All data downloaded from dbGaP for the closed project must be deleted upon project close-out. Only encrypted copies of the minimum data necessary to comply with your institutional scientific data retention policy should be retained. Any copies of data retained must be deleted at the appropriate time and per your institutional records retention policy timelines; this should include any data on central servers or their back-up systems, laboratory computers, or staff/student personal computers.

# Shred hard copies and CD ROMs or other non-reusable physical media.

# Delete electronic files securely.

# At minimum, delete the files and then empty your recycle bin.

# Optimally, use a secure method, e.g., an electronic “file shredder” program that performs a permanent delete and overwrite.

# Most institutions retain backup tapes for a period of time. Ensure those backups are reused (data deleted) and any archive copies are also destroyed.

# Refer to the Guidelines for Information Media Sanitization (http://csrc.nist.gov/publications/nistpubs/800-88/NISTSP800-88\_rev1.pdf) for further details on data disposal

# Additional Resources for testing and best practices: The Center for Internet Security

# CIS is the only distributor of consensus best practice standards for security configuration. The Benchmarks are widely accepted by U.S. government agencies for FISMA compliance, and by auditors for compliance with the ISO standard as well as GLB, SOx, HIPAA, FERPA and other regulatory requirements for information security. End user organizations that build their configuration policies based on the consensus benchmarks cannot acquire them elsewhere.

# http://www.cisecurity.org/.

# Appendix C – Has checklists based on CIS best practices, customized for dbGaP data use. Content of this document has been adapted from CIT/NIH, NIST and CIS.

# 

# APPENDIX C. BEST PRACTICE SECURITY

# REQUIREMENTS FOR dbGaP DATA RECIPIENTS

# Preface

# This appendix has been adapted from the HHS IT Security program for minimal security standards and the Center for Internet Security, and adapted as “Best Practices” for dbGaP. This appendix is intended for the engineering team that will configure local computer system that store and use dbGaP data. There are two parts: general security and a digest of the NIST recommended steps to delete data and scrub devices.

# Introduction

# The dbGaP Best Practices Guidelines checklists were created to provide guidance and expectation on how to treat the controlled access data received from dbGaP.

# Purpose

# The purpose of this appendix is to provide minimum configuration standards for recipients of data from dbGaP. Adhering to these procedures will provide a baseline level of security, ensuring that minimum standards or greater are implemented to secure the confidentiality, integrity, and availability of data resources. If institutional IT policies are more restrictive, then they should apply.

# Background

# Minimum security configuration standards help to ensure sound control of each system. Adhering to minimum standards helps to mitigate risks associated with implementing applications and software by providing a solid foundation to track changes, the differences between versions, and new components as they are installed. System and application default settings are not optimal from a security perspective. Using default settings increases the risk of exploitation. These risks are mitigated through the use of minimum security configuration standards. These standards are from CIS checklists and are cross mapped to NIST Recommended Security Controls for Federal Information Systems 800-53 Rev. 4.

# For additional details, please refer to the benchmarks and audit tools at http://cisecurity.org/

# 

# Microsoft also has security tools such as the Security Compliance Manager that will check and provide feedback on your system’s security

# Windows: XP/2003/2008/2012/Vista/7/8 – Windows Configuration Guide

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Windows Configuration Guide - If action not completed, add comment with explanation** | | | | | |
| **Category** | **800-53** | **800-53 Map** | **Action** | **Completed** | **Comments** |
| Access Controls | Access Enforcement | AC-3 | Only allow Server Administrators to Schedule Tasks |  |  |
| Access Controls | Access Enforcement | AC-3 | Do Not Allow Automatic Administrative Logon |  |  |
| Access Controls | Access Enforcement | AC-3 | Configure all disk volumes to use the NTFS file system |  |  |
| Access Controls | Access Enforcement | AC-3 | Set Unsigned Driver Installation Behavior To "Warn but allow installation" or "Do not allow installation" |  |  |
| Accounts | Account Management | AC-2 | Rename Administrator Account |  |  |
| Accounts | Account Management | AC-2 | Disable or delete unused user accounts |  |  |
| Access Controls | Access Enforcement | AC-3 | Enable account lockout after specific length of time |  |  |
| Accounts | Account Management | AC-2 | Rename and disable the Guest Account |  |  |
| Accounts | User Identification and  Authentication | AC-3 AC-7 IA-2  IA-5 | Configure the system per 800-53 Account Policy Control Requirements |  |  |
| Accounts | Account Management | AC-2 | Do not allow anonymous enumeration of SAM accounts |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Windows Configuration Guide - If action not completed, add comment with explanation** | | | | | |
| **Category** | **800-53** | **800-53 Map** | **Action** | **Completed** | **Comments** |
| Accounts | Account Management | AC-2 | Do not allow anonymous enumeration of SAM accounts and shares |  |  |
| Accounts | Account Management | AC-2 | Disable anonymous SID/Name translation |  |  |
| Accounts | User Identification  and Authentication | AC-3 AC-7  IA-2 IA-5 | All passwords should be strong passwords (see minimum password requirements), and account names longer than 6 characters |  |  |
| Accounts | Account Management | AC-2 | Limit local account use of blank passwords to console logon only |  |  |
| Additional Security Settings | Configuration Management | CM-2 CM-6 CM-9  SI-3 | Enable Microsoft Security Essentials or other Antivirus/Anti-malware software, and ensure software is up to date |  |  |
|  |  |  |  |  |  |
| Additional Security Settings | Configuration Management | CM-2 CM-6 CM-9 SI-3 | Keep your anti-virus software up-to-date and configure anti-virus software to check for updates daily |  |  |
| Device | Session Lock | AC- 11 | Disable allowing users undock without having to log on |  |  |
| Logon | User Identification and Authentication | IA-2 | Configure the system to display a warning banner. |  |  |
| Logon | User | IA-2 | Do Not Allow System to be Shut Down Without |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Windows Configuration Guide - If action not completed, add comment with explanation** | | | | | |
| **Category** | **800-53** | **800-53 Map** | **Action** | **Completed** | **Comments** |
|  | Identification and Authentication |  | Having to Log On |  |  |
| Logon | User Identification and Authentication | IA-2 | Enable CTRL+ALT+Delete Requirement for Logon |  |  |
| Media | Remote Access | AC- 17 | Restrict CD-ROM Access to Locally Logged-On User Only |  |  |
| Media | Remote Access | AC- 17 | Restrict Floppy Access to Locally Logged-On User Only |  |  |
| Network Access | Account Management | AC-2 | Disable letting Everyone permissions apply to anonymous users |  |  |
| Network Access | Remote Access | AC- 17 | Digitally Encrypt Secure Channel Data |  |  |
| Network Access | Remote Access | AC- 17 | Digitally Sign Client Communication |  |  |
| Network Access | Remote Access | AC- 17 | Digitally Sign Server Communication |  |  |
| Network Access | Remote Access | AC- 17 | Require Strong (Windows 2000 or later) Session Key |  |  |
| Network Access | Remote Access | AC- 17 | Disable Sending Unencrypted Password to Connect to Third-Party SMB Servers |  |  |
| Network Access | Remote Access | AC- 17 | Restrict anonymous access to Named Pipes and Shares |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Windows Configuration Guide - If action not completed, add comment with explanation** | | | | | |
| **Category** | **800-53** | **800-53 Map** | **Action** | **Completed** | **Comments** |
| Network Access | Remote Access | AC- 17 | Configure system so that no shares can be accessed anonymously |  |  |
| Network Access | Transmission Integrity | SC-8 | Do not allow storage of credentials or .NET passports for network authentication |  |  |
| Network Security | Information Remnants | SC-4 | Do not store LAN Manager password hash value on next password change |  |  |
| Network Security | User Identification and Authentication | IA-2 | Configure LAN Manager Authentication Level to "Send NTLMv2 response only\refuse LM" |  |  |
| Password Management | Access Enforcement | AC-3 | Do Not Store Passwords Using Reversible Encryption |  |  |
| Password Management | Authenticator Management | IA-5 | Disable System Maintenance of Computer Account Password (Domain Controllers) |  |  |
| Patches | Flaw Remediation | SI-2 | Apply critical Operating System security patches |  |  |
| Patches | Flaw Remediation | SI-2 | Enable automatic notification of patches availability |  |  |
| Patches | Flaw Remediation | SI-2 | Install the latest service packs and hotfixes from Microsoft |  |  |
| Patches | Flaw Remediation | SI-2 | Ensure That Before the System is Loaded Onto an Operational Network, Security Patches, Service Packs, And Hot Fixes are all Tested |  |  |
| Permissions | Access Enforcement | AC-3 | Configure the system to provide least access to shared folders |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Windows Configuration Guide - If action not completed, add comment with explanation** | | | | | |
| **Category** | **800-53** | **800-53 Map** | **Action** | **Completed** | **Comments** |
| Registry Permission | Least Functionality | CM-7 | Disable CD Autorun |  |  |
| Registry Permission | User Identification  and  Authentication | IA-2 | Disable Automatic Logon |  |  |
| Service | Least Functionality | CM-7 | Configure permissions for the following services to give Administrators 'Full Control' and the System 'Read' and 'Start, Stop, and Pause.'  Alerter (Alerter)  Client Service for NetWare (NWCWorkstation) Clipbook (ClipSrv)  Fax Service (Fax)  File Replication (NtFrs)  File Server for Macintosh (MacFile) FTP Publishing Service (MSFtpsvc) Help and Support (helpsvc)  HTTP SSL (HTTPFilter)  IIS Admin Service (IISADMIN) Indexing Service (cisvc)  License Logging Service (LicenseService) Messenger (Messenger)  Microsoft POP3 Service  NetMeeting Remote Desktop Sharing (mnmsrvc) |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Windows Configuration Guide - If action not completed, add comment with explanation** | | | | | |
| **Category** | **800-53** | **800-53 Map** | **Action** | **Completed** | **Comments** |
| Service | Least Functionality | CM-7 | Network Connections  Network News Transport Protocol (NNTP) (NntpSvc)  Print Server for Macintosh (MacPrint)  Print Spooler (Spooler)  Remote Access Auto Connection Manager (RasAuto)  Remote Access Connection Manager (RasMan)  Remote Administration Service Remote Desktop Help Session Manager (RDSessMgr)  Remote Installation (BINLSVC) Remote Procedure Call (RPC) Locator  (RpcLocator)  Remote Registry Service (RemoteRegistry) Remote Server Manager (AppMgr)  Remote Server Monitor (Appmon)  Remote Storage Notification (Remote\_Storage\_User\_Link) Remote Storage Server (Remote\_Storage\_Server)  Simple Mail Transfer Protocol (SMTP) (SMTPSVC)  SNMP Service (SNMP)  SNMP Trap Service (SNMPTRAP) Telephony (TapiSrv)  Telnet (TlntSvr)  Terminal Services (TermService) Trivial FTP Daemon (tftpd) Wireless Configuration (WZCSVC)  World Wide Web Publishing Services (W3SVC) |  |  |

|  |
| --- |
|  |
| **Category** | | **800-53** | **800-53 Map** | **Action** | **Completed** | **Comments** |
| Service | | Least Functionality | CM-7 | Review all services for proper configuration and disable or uninstall unneeded services |  |  |
| Registry Permission | | Least Functionality | CM-7 | Remove administrative shares on servers |  |  |
| User Rights | | Access Enforcement | AC-3 AU-8  AU-9 | Audit user rights assignments to ensure they are appropriately applied |  |  |
| Auditing and Account Policies | | Auditing and Accountability | AU-1 AU-3  AU-6 | Configure audit policies |  |  |
| Auditing and Account Policies | | Auditing and Accountability | AU-2 AU-3  AU-6 | Configure event log |  |  |
| Additional | | Boundary | AC-3 | Make sure your Windows Firewall is on (use |  |  |
| AC-4 |
| Security Settings | | Protection | CA-3 | Action Center) |
| SC-7 |
| Additional | | Boundary | AC-3 | Use the Internet Connection Firewall or other |  |  |
| AC-4 |
| Security Settings | | Protection | CA-3 | methods to limit connections to the server |
| SC-7 |
| Storage and Backup | | Contingency Planning | CP-6 CP-9 | Backup your files and settings regularly |  |  |

**MacOS X/Linux Variants**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **MacOS/Linux Variants Configuration Guide - If action not completed, add comment with explanation** | | | | | |
| **Category** | **800-53** |  | **800-53 Map** | **Action** | **Completed** |
| Accounts / Access | Account Management | AC-2 | No '.' (current working directory) or group/world writable files exist in root's $PATH. |  |  |
| Accounts / Access | Account Management | AC-2 | Install TCP Wrappers |  |  |
| Accounts / Access | Account Management | AC-2 | Remove user .netrc files |  |  |
| Accounts / Access | Account Management | AC-2 | Set "mesg n" as default for all users |  |  |
| Accounts / Access | Account Management | AC-2 | Set default group for root account |  |  |
| Accounts / Access | Account Management | AC-2 | Verify that no UID 0 accounts exist other than root |  |  |
| Accounts / Access | Account Management | AC-2 | Normal use login as user not as an administrator |  |  |
| Account / Access | Account Management | AC-2 | Disable or delete unused users |  |  |
| Accounts / Access | Access Enforcement | AC-3 | Set project directories to be as restrictive as possible to the research group |  |  |
| Accounts / Access | Access Enforcement | AC-3 | Set Account Expiration Parameters On Active Accounts |  |  |
| Accounts / Access | Access Enforcement | AC-3 | Require Authentication For Single-User Mode |  |  |
| Accounts / Access | Access  Enforcement | AC-3 | Remove rhosts support in pam |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **MacOS/Linux Variants Configuration Guide - If action not completed, add comment with explanation** | | | | | |
| **Category** | **800-53** |  | **800-53 Map** | **Action** | **Completed** |
| Accounts / Access | Access Enforcement | AC-3 | Remove empty crontab files and restrict file permissions to authorized users |  |  |
| Accounts / Access | Access Enforcement | AC-3 | Restrict at/cron to authorized users |  |  |
| Accounts / Access | Access Enforcement | AC-3 | Restrict root logins to system console or ssh on local network |  |  |
| Accounts / Access | Access Enforcement | AC-3 | Set LILO/GRUB Password if possible, or set password before boot |  |  |
| Accounts / Access | System Use Notification | AC-8 | Set a warning banner for console and GUI based logins. |  |  |
| Auditing | Auditable Events | AU- 2 | Enable system accounting (Install the sysstat package if needed). |  |  |
| Installation / Patches | Transmission Integrity | SC-8 | Utilize Secure Shell (SSH) for remote logins and file transfers. |  |  |
| Patches, | Flaw Remediation | SI-2 | Apply critical Operating System security patches |  |  |
| Misc / Tuning | Information Flow Enforcement | AC-4 | Deny all network access to the system via hosts.deny; Explicitly allow network connections, either all services selected ones,  from the local network and selected hosts via hosts.allow |  |  |
| Misc / Tuning | Information Flow  Enforcement | AC-4 | Add ' nosuid' and ' nodev' Option For Removable Media In /etc/fstab |  |  |
| Auditing | Protection of Audit Information | AU- 9 | Unless the host is functioning as a syslog server, prevent the system from accepting syslog messages from the network. |  |  |
| **MacOS/Linux Variants Configuration Guide - If action not completed, add comment with explanation** | | | | | |
| **Category** | **800-53** |  | **800-53 Map** | **Action** | **Completed** |
| Auditing | Protection of Audit Information | AU- 9 | Unless the host is functioning as a syslog server, prevent the system from accepting syslog messages from the network. |  |  |
| Misc / Tuning | Least Functionality | CM- 7 | Set default UMASK for users, directories, and files to meet the needs of the system |  |  |
| Misc / Tuning | Least Functionality | CM- 7 | Disable Core Dumps |  |  |
| Services | Least Functionality | CM- 7 | Disable xinetd if none of its services are used |  |  |
| Services | Least Functionality | CM- 7 | Disable Sendmail and other inbound mail daemons |  |  |
| Services | Least Functionality | CM- 7 | Disable standard boot services that do not support the role of the system |  |  |
| Services | Least Functionality | CM- 7 | Turn off standard services except those needed for the system’s role. |  |  |
| Logon | User Identification and  Authenticatio n | IA-2 | Configure the system to display a warning banner. |  |  |
| Accounts / Access | Authenticator Management | IA-5 | No “+" entries should exist in /etc/passwd or  /etc/group. |  |  |
| Installation / Patches | Transmission Integrity | SC-8 | Utilize Secure Shell (SSH) for remote logins and file transfers. |  |  |
| Patches, | Flaw  Remediation | SI-2 | Apply critical Operating System security patches |  |  |

For Mac Laptops, an encryption tool such as FileVault should be used to protect all controlled access data

1. If contractor services are to be utilized, the principal investigator (PI) requesting the data must provide a brief description of the services that the contractor will perform for the PI (e.g., data cleaning services) in Section B of the WIHS Concept Sheet. Additionally, the Lead Investigator section (i.e., Section A5) of the WIHS Concept Sheet must include the name of the contractor’s employee(s) who will conduct the work. These requirements apply whether the contractor carries out the work at the PI’s facility or at the contractor’s facility. In addition, the PI is expected to include in any contract agreement requirements to ensure that any of the contractor’s employees who have access to the data adhere to the WIHS Genomics Data Use Contract, the WIHS Genomics Data Code of Conduct, and the WIHS Data Security Best Practices Requirements. [↑](#footnote-ref-1)