NBER's data security policies and procedures for hosting confidential CMS data.

Revised: April, 2023

1) **Computing Environment:**

NBER's computing facilities are based on centralized UNIX and Linux servers with network attached storage servers. All of the computer systems with hard drives and other "online" data storage facilities containing restricted data will be kept in our centralized locked server rooms. Physical access to the locked server room is provided via electronic keypad. Entries to the server room are logged on NBER Office Manager's computer, who also manages access to the building premises. Access card/keys to the server rooms are programmed and provided only to NBER IT staff. Should there be a need for an outside IT Consultant, entry into the room will be in the presence of authorized NBER personnel. Except when authorized NBER personnel are physically present in these server rooms, they will be kept locked at all times. Further, the main entrance to the NBER offices is on the 3rd floor of the building, where a receptionist is present. All other entrances to the NBER premises have electronic locks and are monitored with closed-circuit televisions.

Remote login access to authorized users is provided via secure shell or xrdp service, and all of the processing, computation, and analytic work using sensitive data is to take place on these servers and the storage devices. Statistical software such as SAS and Stata, and computational packages such as R, Python, Octave etc. are installed on these servers for computing. So all the analyses including the data can remain within the confines of these servers located in the locked rooms, even though the users mostly connect remotely from locations outside the physical location of these servers. Users pledge not to take out any output containing any identifiable data. All physical (removable) disks and tapes containing restricted data at the NBER are kept in locked cabinets in these locked server rooms. On-site backups are also kept in these locked server rooms. Offsite backups are encrypted and will be sent electronically, in encrypted form, to a secure co-location facility. This co-location facility(currently Cyxtera Co-location, Waltham, MA) is a reputed highly secure Datacenter that includes 24/7 monitoring, with individually locked cages. Entrance to the Datacenter requires a physical/Badge ID check at the Security Desk. Entrance to the work area requires a 2FA process, first with a pre-issued access card and followed by bio-metric authentication(finger print or hand scan) of authorized users before entry to the premises is permitted. The appropriate NBER IT staff will be pre-registered and vetted by the Datacenter for authorization to receive access cards. The assigned cage for NBER, in which NBER hosts its backup servers, is secured by a number lock. Unless requested by NBER authorized personnel, no one will enter the cage including the Datacenter Administrators. The data transmitted electronically will be encrypted from end to end and will remain encrypted, at rest, on the server at the co-location facility. The encryption will be using FIPS 140-2 compliant AES-256 or more technology.

Several additional measures have been put in place on the LAN. Servers for computational use are separated out from systems for general activities such as webserver, mailserver etc. Any programs that transmit information in clear (unencrypted) text, such as "telnet" and "ftp" are disabled on these systems. Research computers are booted “diskless” from a central storage server. They are deliberately booted in read-only mode so even under a breach of an individual research server the central core is not compromised. A select set of “trusted” servers are set up that have read & write access to the diskless volumes. This allows for software installations, OS and patch updates to be performed in a controlled fashion. Login to the “trusted” servers is restricted to only a few NBER IT accounts, and 2FA is enforced. User account passwords are vetted using “pam_passwdqc” utility, and strong passwords are enforced. A vulnerability scanner (Nessus) runs once a week on the entire address space on the LAN. A honeypot computer has been setup to bait and block intrusion attempts.

Many additional security procedures are applied to communications between the in house NBER computer systems and computers outside the walls of the NBER offices. At our gateway we have an application firewall(currently Fortinet’s Fortiguard appliance) that provides additional protection. We subscribe to real-time virus checking from the firewall vendor. All web and email traffic are subject to this virus checking therefore the office LAN is protected from virus and malware at the gateway. The internal LAN is also segmented into several VLANs and the VLAN for the network hosting confidential data are on a private address space protected by a VPN, details of which are elaborated later in the document. Firewall rules, and Access controls (acls) are used extensively to granularly mange access to services and protocols between the VLANs as well as to the outside WAN. By default, no ports except for ssh and xrdp ports are open to the outside world. Specific ports based on functionality of a server are open for that server. For example, smtp & (s)pop/(s)imap ports are open only for the mailserver(s), while http/https ports are open only to the webserver(s). As mentioned before, research computing servers and storage appliances are not part of such roles and therefore they can be accessed only via the ssh protocol or Remote Desktop protocol. Further the xrdp service is configured at the server end where the “cut and paste” or transfer features are disabled.

VPN ENVIRONMENT: The servers hosting confidential data covered by this agreement are placed inside a private LAN(a separate VLAN) isolated from the general office LAN behind a NATting firewall that requires an encrypted VPN connection. These servers inside the private LAN have an IP address in RFC1918 space (a private address space not forwarded to the outside world). Since connection to the private LAN requires an established VPN connection, all transactions to and from the private LAN are tunneled via the VPN in encrypted form. Additionally, each VPN connection is authenticated using a two factor authentication system. The first authentication is based on a user’s UNIX username/password (strong passwords are required vetted via pam_passwdqc) via a radius server. The 2nd factor is currently administered via DUO Security service. After the password is authenticated, the user re-confirms themselves either via a DUO Mobile push acknowledgement or by receiving a phone call/SMS to a

registered user phone number that the user acknowledges by answering and pressing # on their phone or by providing a OTPW (one time password) from a registered hardware token associated with the user. The IT staff at NBER will register each phone number for each user authorized to use these systems. This mechanism will act as a protection against all password compromises including a remote client that could be virus/malware infected, keystroke logging security breaches, shoulder surfing, phishing, etc. The VPN is configured to lock an account after three consecutive failed log-in attempts. DUO service is also configured to disable a registered account when the user has not logged in over the last six months. The firewall is configured to allow certain ports/protocols between the VPN and the non-VPN environment. These are primarily to assist in system administration and monitoring (for example for monitoring via BigBrother, NAGIOS and nrpe, or for logging to central syslog service, or for synchronizing the operating systems under our diskless architecture).

Unix user groups/ACLs will control access to directories within these secure servers. Unix groups will be made based on DUA/Project and researchers are appropriately associated with their respective groups. The data management team has its own group. Each DUA/Project will be assigned specified directory space where researchers on the DUA/Project must keep all their files and conduct their analyses.

2) Transmission of Electronic Confidential Data:

NBER will receive confidential data from the source agency in encrypted form. The encryption key shall be sent separately from the encrypted data. Confidential data will not be shared with any other institution or facility unless proper approval has been obtained from the data provider.

3) Storage of Confidential Data:

All “online” confidential data will be stored within NBER’s centralized computing environment as described in item 1.

Original data received from CMS after being uploaded to these secure servers will be kept in locked cabinets/safe within these locked server rooms or a locked safe. Only authorized personnel will have access to the cabinets. If the data is received via a secure transfer or download facility(e.g. sftp/scp) then only the on-site copy on the servers and their backups will be kept.

On-site backup of confidential data will also be stored on separate servers within the locked server rooms. Only systems administrators will have access to these backup servers.

Off-site backups are encrypted and will be sent electronically, in encrypted form, to a secure co-location facility. This co-location facility(currently Cyxtera Co-location, Waltham, MA) is a reputed highly secure Datacenter that includes 24/7 monitoring, with individually locked cages. Entrance to the Datacenter requires a physical/Badge ID check at the Security Desk. Entrance to the work area requires a 2FA process, first with a pre-issued access card and followed by biometric authentication(finger print or hand scan) of authorized users before entry to the premises is permitted. The appropriate NBER IT staff will be pre-registered and vetted by the Datacenter for authorization to receive access cards. The assigned cage for NBER, in which NBER hosts its backup servers, is secured by a number lock. Unless requested by NBER authorized personnel, no one will enter the cage including the Datacenter Administrators. The data transmitted electronically will be encrypted from end to end and will remain encrypted, at rest, on the server at the co-location facility. The encryption will be using FIPS 140-2 compliant AES-256 or more technology.

4) **Authorization Procedure:**

Each project will apply with CMS and receive appropriate approvals or Data Use Agreements. A copy of the agreement will be submitted to the Data Custodian/Data Manager and NBER IT. The project will also undergo NBER IRB approval.

Prior to receiving authorization to use restricted data at the NBER, potential new researchers or staff will take a Human Subjects Protection training course. The researcher will then be briefed on the NBER's computing facilities and procedures on maintaining data security in effect at the NBER, and the penalties for failing to comply with these procedures. After completing this training session the researcher/staff will sign a pledge of data confidentiality with NBER.

5) **User Account setup:** After receiving appropriate authorization, a user will be assigned a login userid. Not only is each user assigned a unique account, a DUA specific account is also issued. The account credentials will be required to create a strong password that will be vetted by "pam_passwdqc" utility. Users would have to login separately using their DUA specific account to access data authorized under that DUA and to the project computing space assigned to them. Only our data management team may be exempt from this requirement for administrative purposes. Each project/DUA is associated with a unix group that governs access to the project's authorized data as well as the assigned work space on the servers. Accounts are created from a centralized system and the accounts database is backed up nightly allowing us to review changes if required. Monitoring and review of accounts are performed in a couple of ways. Firstly, each researcher account needs to be sponsored by NBER affiliated faculty. Annually, the faculty member is sent an automatic reminder of the list of accounts they sponsor and whether they would continue to sponsor the accounts. If the faculty member "de-sponsors" an account, the account is deactivated by an automatic "cron" job that is run monthly. Secondly, nearing

each DUA's annual expiration date, the NBER's data and DUA management staff presents a list of current members of a DUA to the DUA Requester/PI and seek a verification of the list of users before submitting a request for an extension should there be a need for an extension. The NBER staff follows up by requiring each member who needs continued access to take our annual CMS Data users training. All members the DUA Requester/PI indicates as to be removed are removed from access. At the advice of the DUA Requester/PI, users can be added or de-activated out of the above cycles as well, including on an emergency basis. DUO security, which is used for 2FA of the VPN, is configured to deactivate the account if a user has not logged on for nine months. These are some of the ways in which user accounts are managed. Temporary accounts or emergency accounts are not registered for VPN access so they do not have access to these systems.

6) **Access to Data:**

After receiving appropriate authorization, a user will then be assigned to an appropriate user group(s) based the project DUA(s). Login access to the secure servers will then be enabled for the user. Procedures to login will follow the guidelines described in item 1 (Computing Environment). Unix user groups along with ACL/permission settings will control access to these directories.

All work using the confidential data will be performed on the secure servers. Each project/DUA will be allotted specific directory folders where all the work for the DUA must be conducted. CMS datasets are large and multiple projects obtain authorization to use the same data. In order to have efficient use of resources of potentially enormous datasets, datasets common to multiple projects will be kept in a common pool of directories with read-only access. However, access to individual files within these common pool of directories will be granted for each project based on data authorized under the project's DUA. The default setting is a denied access for all users but for the data management team, and ACLs will be set to grant access on a per DUA and per dataset basis.

7) **Output Review:**

Users are instructed to have their results, summary statistics, and output reviewed by NBER's Data Custodians/Managers/IT personnel. Only analytical results and aggregates not containing any identifiable data can be taken out. The NBER personnel, to the best of their abilities, will review the requested output for any inadvertent disclosure of personally identifiable information before they take it out of this VLAN.

8) Data Destruction and Disposal:

The Principal Investigator of a DUA will inform the Data Custodian/NBER IT staff of closure of a project. At the end of a project (as informed by the Researcher) the restricted data will be deleted from the servers using a program like "shred" or equivalent. Shred will over-write all deleted sectors of the data so the data is not recoverable. Any retired data disks will be first "zeroed out". Any CD/DVD ROMs will be destroyed and magnetic tapes will be degaussed with an onsite degausser.

After the data has been purged, the Requester or Data Custodian will then complete the Certification of Data destruction form at http://www.cms.gov/Medicare/CMS-Forms/CMS-Forms/downloads//cms10252.pdf and submit it to CMS.

9) Termination of access:

The PI will inform the Data Custodian/NBER IT staff of any termination of participation of a researcher in their project DUA. The NBER IT staff will then remove the user from the unix group membership for those projects, and if necessary remove from login access to the servers that are dedicated for these restricted data analyses or the private LAN itself. Correspondingly, the IT Director will handle access changes for any other member of the NBER IT staff, and the President of NBER will initiate and supervise access changes for the IT Director, should such a need arise.

As a secondary step, annually, the Data Custodian/NBER IT staff will present the PI with a list of authorized users and seek an update. Accordingly, access to individuals will be revised.

9) Privacy Breach

he PI has the main responsibility for informing CMS by phone: 410-786-2580 or 1-800-562-

1963 or e-mail: CMS_IT_Service_Desk@cms.hhs.gov of any suspected incidents wherein the security and

privacy of the CMS data may have been compromised. The PI will contact CMS via phone or email in

compliance with CMS' Privacy Breach policy at http://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/Privacy/Privacy_Data_Breach.html. The IT Administrator(s) will

inform the PI if they identify any incidence of breach on the computing environment.