> **SJÖGREN'S INTERNATIONAL COLLABORATIVE CLINICAL ALLIANCE (SICCA)**
> **BIOREPOSITORY AND DATA REGISTRY**
> **DATA MANAGEMENT PLAN**

1. Database Overview

The SICCA database includes the Clinical Database, the Biorepository Database, and the DNA database which are housed in separate locations.

SICCA Registry Clinical Database

A computerized, secure SAS database system contains all clinical data derived from SICCA study participants (Clinical Database). The content of the Clinical Database consists of basic demographic information of participants including gender, age at time of enrollment in the baseline SICCA study visit, ethnicity, and country of residence, as well as research records from baseline and follow-up study visits (antibody tests, results of clinical assessments, etc). For a list of content see Clinical Database Data Dictionary (Attachment 1). The Clinical Database does **not** contain any Protected Health Information (PHI*) as described and defined by HIPAA standards. As such, any data sets released via this database system meet the HIPAA definition of **De-Identified Data**.

*The Clinical Database does **not** contain any of the following elements:

- Names
- All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements will be aggregated into a single category of age 90 or older;
- All geographical subdivisions smaller than a State, including street address, city, county, precinct, zip code, and their equivalent geocodes
- Telephone numbers
- Fax numbers
- Electronic mail addresses
- Social Security numbers
- Medical record numbers
- Health plan beneficiary numbers
- Account numbers
- Certificate/license numbers
- Vehicle identifiers and serial numbers, including license plate numbers
- Device identifiers and serial numbers
- Web Universal Resource Locators (URLs)
- Internet Protocol (IP) address numbers
- Biometric identifiers, including finger and voice prints
- Full face photographic images and any other comparable images
- Any other unique identifying number, characteristic, or code (note this does not include the randomly generated study id used by the researchers to organize the data)

Additionally, prior to any data set release as part of the dissemination plan, the data will

be screened for records with particularly unique characteristics, such that even in the absence of PHI the individuals might be identifiable –i.e. an African-American of a certain gender, age, and with a specific illness living in Denmark. Such "rare" records will be grouped into a more general category –i.e. "other" for ethnicity –to further mask their identity.

SICCA Biorepository Database
The SICCA biorepository database is located at the AIDS Specimen Bank (ASB) on the Parnassus Heights campus. Biospecimens managed by the ASB include plasma serum, PBMCs, tears, conjunctival RNA imprints, whole saliva, parotid saliva, frozen minor salivary glands, paraffin blocks of minor salivary gland biopsies and both H&E- stained and unstained slides. The ASB database tracks the exact locations of SICCA biospecimens within freezer boxes, number of aliquots, where an aliquot has been shipped, and number of remaining aliquots. Each location combination (e.g. freezer, rack, box, row and column) is uniquely identified within the ASB database system (following "Best Practice" as cited by NCI and the International Society for Biological and Environmental Repositories or ISBER). SICCA specimen aliquots are identified by the ASB accession number which is distinct from the SICCA participant number. There are no personal identifiers on the labels. The database is backed up daily at the ASB and on the School of Medicine's server weekly, which is then moved to an off-site data storage facility, Iron Mountain.

SICCA DNA Database
The UC Berkeley Genetics and Genomics Laboratory (UCB Lab) Database tracks all steps of processing and storage of DNA extracted from SICCA participants and from participant relatives. Data related to the samples such as collection date, received date, processed date, concentration, initial yield, volume, and purity are recorded in the database. Aliquots and corresponding concentration and volume are also recorded. The 2D barcode associated with all DNA sample tubes allows easy verification of sample and sample location when retrieving/checking as needed or during annual checkups.

2. SICCA Website and Dissemination
Plan Website
The purpose of the SICCA Website is to introduce potential requestors to this project and allow them to explore a catalog of available clinical data and specimen types so that they may formulate a research proposal. Potential requestors will create a User Account on this website and once they have submitted a signed Data Use Agreement electronically, will be able to use the Search/Browse and Data Explorer functions to identify the criteria and variables by which they would like to request a clinical dataset. A catalog of available biospecimen types will also be kept on the website for potential requestors to consult. The purpose of the Data Explorer feature is to provide potential requestors with the opportunity to get familiar with the clinical data and thus to refine their research question. Using the Data Explorer function, potential requestors will be able to query the clinical database, but they will not have access to individual-level records, and will not be able to download any data. Requestors will also be able to browse data collection forms, annotated to include variable names, and data dictionaries of any computed or derived variables. They will then be able to submit a

Letter of Intent (LOI) via the website followed by a full application which will go through the review process previously described in our technical proposal.

Dissemination Plan
Once an investigator's request has been approved and material transfer agreements have been signed and approved, the specified clinical data and/or biospecimens will be released in the following manner: Specimens will be sent from the ASB or the UCB Lab to a requestor with their Specimen ID. The requestor will run their assays and generate a test result data set, which they will then submit back to the SICCA Data Manager. The data center will match the test result data set to their corresponding clinical data which will then be made available to the requestor as a merged dataset, customized to the needs of their proposal. This customized dataset will be downloadable from the website using a release ID and a password protected protocol. The requestor will not see the conversion key between the 2 numbers (release ID and Specimen ID) nor will they ever see the SICCA Participant ID.

3. Disaster Recovery Plan
As previously described in our technical proposal, we are providing the following summary of our Disaster Recovery plan:

Database Backup/Testing
We conduct tests yearly along with the audits. Databases are tested for integrity every night before backup. Such testing is a standard part of any database backup and is implied. Databases are backed up three times daily including one time offsite. The actual backup data itself is tested nightly for integrity, as it is built into the system. Also, "user restores" are performed more than a few times a week as part of our testing. All restores are logged and tracked.

Security at Backup Site
The backup site is a third party co-location facility owned by Virtustream whose security and compliance highlights include SSAE16, ISAE34021, PCI-DSS, FISMA2 and HIPAA/HITECH4 compliance. As well as 24 x 7 on-site security including closed video surveillance, multi-level authentication for access control including biometrics, keycards and controlled access at all times.

Emergency Backup Systems
1) For AIDS Specimen Bank (ASB) emergency back-up systems: The ASB protocol for addressing out of work hours emergencies, described in detail in the Technical Proposal, is as follows. One staff member is assigned a pager that receives messages from the alarm system. This staff member is not permitted to be on vacation when the pager is assigned. The remaining ASB staff members are contacted automatically by email and text messages by the alarm system. The alarm system is programmed to contact the pager first, then ASB staff and Directors, then Oyster Point (OP, were the Freezer Farm is located) staff. The person assigned the pager is also responsible for contacting OP staff and the co-director and data manager of ASB of the situation. OP staff address freezer or power failure at the Freezer Farm. ASB staff are not allowed to enter the Freezer Farm after hours unless accompanied by an OP staff member or campus police. If the freezer or power outage occurs on the Parnassus

campus, the person wearing the pager is responsible for going to campus and addressing the issue. The ASB lab maintains an empty freezer for emergencies. It is ASB policy that all staff are responsible for being part of this process except for when on sick leave or vacation. It is ASB policy that the first person who addresses a freezer emergency must acknowledge this on the alarm's web page. Thus, the date and time is stamped, documenting when the situation was addressed. It is procedure that an 'alarm' situation be addressed within one hour of receiving the message. Again, the staff member assigned the pager must be the first to address the issue.

2) UC Berkeley Genetics and Genomics Laboratory emergency back-up systems:

The laboratory procedure manuals at UC Berkeley Genetics and Genomics Laboratory, where all of the DNA samples are stored, are updated annually to reflect changes in protocols and new techniques. Electronic and hard copies are maintained. All equipment (including Refrigerators/Freezers, Centrifuges, other) are monitored by lab personnel. Maintenance, calibration, and repairs are performed as needed by qualified service technicians recommended by the manufacturer or original vendor. The lab refrigerators and freezers are connected to Emergency Power in Stanley Hall. Thus, if a power outage occurs in the building, all refrigerators and freezers are automatically switched to Emergency Power until power is restored. All freezers (where samples are stored) are connected to a central alarm system which provides 24 hour-a-day monitoring and automatic notification to several laboratory staff (via cellphone) in the event of freezer malfunction. The alarm system also notifies staff via cellphone if there is a switch from regular power to Emergency power. There are backup freezers on site available for emergency usage in an event of a freezer malfunction. A plan is in place for emergency transfer of samples. Staff members are also available to respond to any alarm after hours and on weekends.

3) For data server emergency back-up systems:

A technician is on call 24/7. On call tech is required to be reachable. The primary site has no backup power generators, it has only datacenter level UPS for hardware shut down. Backup site has multiple generators and has not lost power in the last ten years. Expected time to respond to alerts: A technician is on call 24/7 and response time varies according to the severity of the issue. Prime production hardware is backed by a two to four hour 24/7/365 parts replacement service. Since we do not control the utility power distribution at the primary site, power will be restored based on PG&E emergency response times.

Selected Metrics:	Down time	Time Period Covered
Core Switch	0 hrs 21 mins	7.6 years (24 hrs PG&E
		Construction at CBL not incl)
C17 Prime SQL Svr	0 hrs 15 mins	
		Construction at CBL not
650SQL1 Prime SQL	10 hrs 28 mins	7.6 years (10 hrs

virtual conversion)

C13 Web Services   0 hrs   13 mins   7.6 years (24 hrs PG&E

Construction at CBL not