

## SECTION 8.0 – TOOLS: THE WEBSITE, THE WAREHOUSE, AND OTHER HELPFUL REPOSITORY TOOLS

### 8.1. Tools for Investigators / Repository Users

- 8.1.1. The Repository Website, Swan Repository.com
- 8.1.2. The Data Warehouse
  - 8.1.2.1. Overview
  - 8.1.2.2. Data Warehouse Content and 4-Step Path
  - 8.1.2.3. Updating the Data Warehouse
- 8.1.3. The Application Submission system
- 8.1.4. Biospecimen Tables
- 8.1.5. Fees
- 8.1.6. Materials for Approved Applications
- 8.1.7. Other Investigator Tools

### 8.2. Tools for Reviewers

The Repository Admin Center Website for Application Reviewers

### 8.3. Internal Tools for Repository Staff

- 8.3.1. Application Management
- 8.3.2. SAS Code Generation from the Data Warehouse

## 8.0 TOOLS: THE WEBSITE, THE WAREHOUSE, AND OTHER HELPFUL REPOSITORY TOOLS

### 8.1. TOOLS FOR INVESTIGATORS/REPOSITORY USERS

#### 8.1.1. The Repository Website: [SwanRepository.com](http://SwanRepository.com)

SWAN Repository provides for effective utilization of its specimens and data through online tools and resources. The SWAN Repository website ([swanrepository.com](http://swanrepository.com)) provides visibility for the Repository and is the primary point of contact for researchers who are interested in using SWAN Repository specimens. The website was designed to create a contemporary and user-friendly hub of information about the Repository and its resources. Several important tools for investigators are located there, including: the Data Warehouse, the automated online application submission system, summaries of available specimens, specimen-associated costs, forms and references for approved investigators, contact information, and answers to frequently asked questions.

#### 8.1.2. The Data Warehouse

##### 8.1.2.1. Overview

The Data Warehouse is a searchable directory or index of SWAN data, and has become a valuable tool in the management and utilization of SWAN resources. With nearly 20,000 variables collected thus far in the SWAN study, finding specific SWAN data can be daunting even for associated investigators. The Repository constructed the Data Warehouse to help make the data more easily navigated. Without allowing direct online access to the data (which is not allowable under current IRB approvals or SWAN study protocols), the Data Warehouse organizes the nearly 20,000 collected and created SWAN

variables into a searchable directory, based on a keyword search engine, and built using the internationally-recognized National Library of Medicine’s MeSH (Medical Subject Headings) hierarchy of medical categories as a guide. Users can choose topics or keywords of interest, then drill down to more specific subtopics (see DWscreenshot1), or search broadly via an open search box (see DWscreenshot2). Users can view all associated variable names and descriptions, response codes, collection methods, assay descriptions, and study years collected, and then select variables, assemble and save them into personal libraries on the site. As studies use Repository specimens to generate more data, this data is returned and also made available through the Data Warehouse inventory and directory.

### Data Warehouse Screenshot1: Search by Keyword

data warehouse REPOSITORY

» LOGGED IN: STEFFENIE MERILLAT » DATA WAREHOUSE » SWAN REPOSITORY

## User libraries: Find Variables by Keyword

Add variables to an existing library, or you can create a new library.

### Step one: Search by keywords

Select keywords and click the  button at the bottom of the page.

**Legend:**

Open and close keyword lists

Select keywords; also selects subordinate keywords

- Demographics/ Participant Description
- BioMeasures
- Genetics
- Reproductive Health and History
- Health History
  - General
  - Cardiovascular
  - Eating disorders
  - Endocrine System
  - Digestive System
  - Nervous System
    - Epilepsy
    - Vasomotor symptoms
    - Symptom sensitivity
    - Pain
    - Migraines

- Hematology

## Data Warehouse Screenshot2: Open Search Box

data warehouse

» LOGGED IN: STEFFENIE MERILLAT » DATA WAREHOUSE » BIOSPECIMENS

### STEP ONE: ENTER YOUR SEARCH TERM

DO YOU THINK YOU KNOW THE NAME OF THE **VARIABLE** YOU'RE LOOKING FOR? ENTER THE NAME BELOW.

SEARCH FOR VARIABLE

WOULD YOU LIKE TO SEARCH THE **DESCRIPTIONS** OF VARIABLES AND KEYWORDS? ENTER YOUR SEARCH TERM(S) BELOW. (RESULTS MAY TAKE 10-20 SECONDS.)

SEARCH FOR DESCRIPTION

### 8.1.2.2. Data Warehouse Content and 4-Step Process

The Data Warehouse describes available data:

- ❖ **SWAN Data** By the end of SWAN's Visit 13 (2013), there were 154 different questionnaires or instruments used to collect or create the nearly 20,000 variables that make up the body of SWAN data. Each of these data variables has been individually linked to keywords in the Data Warehouse. Users can view descriptions of variables, wording of study questions, methodology behind created variables, then create and save lists of variables they are interested in (see 4-Step Screenshot below).
- ❖ **Repository-Generated Data** Data or results generated from SWAN Repository specimens are returned to the Repository within 3 years of close of funding. These new data are then incorporated back into the body of SWAN data, and are included in the Data Warehouse directory system. The origins and methodology behind these data are available as well.

### Data Warehouse Screenshot3: 4-Step Process

**SWAN Resources: Find out what's here**  
Using the Data Warehouse, finding answers to questions about SWAN data can be done quickly and easily.



**STEP 1:** Select Keywords  
Or Enter Search Term



**STEP 2:** Select Visit Years



**STEP 3:** Choose Variables



**STEP 4:** Save List

#### 8.1.2.3. Updating the Data Warehouse

The Data Warehouse database is updated annually, as the SWAN Coordinating Center (CC) releases newly cleaned data from the prior visit. Newly released variables are linked to the keyword structure by Repository staff and added to the database, along with descriptions and other supporting documentation.

### 8.1.3. The Automated Online Application Submission System

Application forms are available to investigators via the Repository website. After analyzing the Data Warehouse, available biospecimen summaries and other online information, investigators can complete and submit online the 2 application forms necessary for requesting Repository resources – the Inquiry Checklist and the full Application (see screenshot below). PDF documents of both forms are available to preview on the website, prior to beginning the submission if desired. (See section 6 for complete application process.)

#### Website screenshot: Online Inquiry Checklist and Application Forms

**Requesting Data & Specimens**

Applying for SWAN Repository resources can now be done online, using these forms:

- STEP 1: Complete the **Initial Inquiry/Checklist** \*.
  - This brief initial assessment allows Repository staff to confirm availability of requested materials and provide a sample size estimate to applicants.
  - [Click here](#) for a preview of the Inquiry Checklist form.
- STEP 2: Complete the full **Application**.
  - Similar to NIH applications, this form includes the major sections: Introduction, Specific Aims and Hypotheses; Background & Significance; Preliminary Studies; and Methods & Materials.
  - The submitted Application will be reviewed by:
    - [SWAN Repository Organization \(SRO\)](#)
    - [SWAN Repository Advisory Group \(SRAG\)](#), if the Application includes requests for DNA samples, genetic data, or Reserved samples
  - [Click here](#) for a preview of the Application form.
- See the [full schematic](#) of Repository application and review processes

### 8.1.4. Biospecimen Tables

The Repository website also provides details about the SWAN specimens available for study. The *Biospecimens* page contains a summary table for each available specimen type (serum, plasma, urine and DNA) showing the number of women with samples available, average number of vials per woman, and average vial volumes. These tables are updated on an annual basis, or subsequent to a large sample pull.

#### Website screenshot: Available Biospecimen Tables

<b>SERUM</b>												
<b>Visits:</b>	<b>Y00</b>	<b>Y01</b>	<b>Y02</b>	<b>Y03</b>	<b>Y04</b>	<b>Y05</b>	<b>Y06</b>	<b>Y07</b>	<b>Y08</b>	<b>Y09</b>	<b>Y10</b>	<b>Y12</b>
<b>Women with Specimens Available</b>	3228	2701	2519	2351	2230	2170	2085	2048	1463	1994	1908	2006
<b>Vials per Woman (Avg)</b>	3.4	8.8	16.4	16.1	14.8	15.2	14.2	15.3	17	16.1	16.7	14.7
<b>Volume per Vial (Avg)</b>	1.05 mL	0.48 mL	0.56 mL	0.55 mL	0.54 mL	0.55 mL	0.55 mL	0.54 mL	0.58 mL	0.56 mL	0.56 mL	0.47 mL

### 8.1.5. Fees

Specimen-related cost tables, which provide current cost-recovery rates for obtaining samples and creating individualized matching datasets, can be viewed on the Repository home page under *Available Biospecimens*.

#### Website screenshot: Link to Fees associated with Obtaining Biospecimens

### The SWAN Repository – Resources available to you

The SWAN Repository is the biospecimen bank of the SWAN study. All stored specimens are from the 3,302 SWAN participants. There are currently nearly 1.8 million samples, collected across the 14 clinic visits. Both SWAN Specimens and Data are available to the scientific community.

- **Available Biospecimens** include Serum, Plasma, and Urine from annual follow-up visits, as well as DNA and Immortalized Cell Lines on approximately 1500 SWAN participants. Also, a subset of 800 women have provided daily urine samples during one menstrual cycle each year until reaching menopause. This collection is known as the Daily Hormone Study (DHS).
  - SWAN Repository specimens have been collected with a high degree of consideration - every step of collection, processing, and storage have been recorded.
  - Some samples, due to scarcity, are held as "reserved" and require a more rigorous review to gain access to. [Click here to see the criteria for Reserved Samples.](#)
  - [See fees associated with obtaining biospecimens.](#)



#### Website screenshot: Current Biospecimen-Related Costs

### SWAN Biospecimen Costs

	Serum, Plasma, Urine	Extracted DNA	Transformed Cell lines
Direct Costs	\$18	\$65	\$148
Indirect Costs	+30%	+30%	+30%

Rates valid through 10/21/2015.

- These prices are based on actual costs accrued in the collection, storage, retrieval and distribution of SWAN samples.
- An official quote will be provided to individual applicants when the final sample selection is defined.

### SWAN Application & Data set fees

	Application Processing Fee	Encrypted Data Set Construction Fee
Direct Costs	\$400	\$950
Indirect Costs	+30%	+30%

Rates valid through 10/21/2015.

### 8.1.6. Materials for Approved Applications

From the website's home page, investigators approved to receive Repository resources can find a number of important forms and references.

**Website screenshot: Link to Materials for Approved Applications**

#### Requesting Data & Specimens

Applying for SWAN Repository resources can now be done online, using these forms:

- STEP 1: Complete the **Initial Inquiry/Checklist** \*.
  - This brief initial assessment allows Repository staff to confirm availability of requested materials and provide a sample size estimate to applicants.
  - [Click here](#) for a preview of the Inquiry Checklist form.
- STEP 2: Complete the full **Application**.
  - Similar to NIH applications, this form includes the major sections: Introduction, Specific Aims and Hypotheses; Background & Significance; Preliminary Studies; and Methods & Materials.
  - The submitted Application will be reviewed by:
    - [SWAN Repository Organization \(SRO\)](#)
    - [SWAN Repository Advisory Group \(SRAG\)](#), if the Application includes requests for DNA samples, genetic data, or Reserved samples
  - [Click here](#) for a preview of the Application form.
- See the [full schematic](#) of Repository application and review processes
- See the [current list](#) of Completed and Approved Repository Projects
- [Materials for approved applications](#)
  - Required forms and Agreements for investigators with approved Repository applications are also available online

These include the Material Transfer Agreement (MTA) and Data Use Agreement (DUA) forms, tutorials for creating variable lists, and instructions for submitting annual progress reports, properly acknowledging SWAN, and returning new data to the Repository.

**Website screenshot: Forms and References available to Approved Investigators**

#### Resources for Approved Applications

Once Repository Applications are fully approved through the Repository review process, investigators will need to submit the following forms:

**Supplemental Form:** Required for all approved biospecimen requests.

[Repository\\_ApplicationSupplementForm.docx](#)

**Material Transfer Agreement (MTA):** Required for all approved biospecimen requests, after funding is secured.

[FORM\\_MTA\\_100110.pdf](#)

**Data Transfer Agreement (DTA):** Required for all approved requests for data only (no specimens).

[DataUseAgreement.pdf](#)

**CSAP Concept Sheet** for new SWAN publications, submitted when approved applications are funded, or used as an application for data (only) requests not seeking new funding.

[CSAP\\_Submission\\_Form.docx \(Microsoft Word\)](#)

#### Additional resources:

##### How to create your datasets

Investigators with approved and funded studies may use the Data Warehouse to browse, compare, and assemble lists of variables that will be used to achieve study aims. ([Click here](#) to assemble lists by keyword or [click here](#) to assemble lists by custom search.) These variable lists may be saved and edited. When complete, the lists will be constructed into cross-sectional SAS datasets at the Repository and sent to the approved study personnel via a secure online file transfer folder (M+Box).

**Annual progress reports**

1-2 page progress reports are due from all approved applicants annually from approval date (date of Approval Letter) until data has been returned to Repository.

**Acknowledgements: For Manuscripts & for Abstracts, Posters, etc** (short version)

[ACKNOWLEDGMENTS\\_forReposWebsite.docx](#) (Microsoft Word)

**How to return data to the Repository; When data is considered for integration with Core SWAN data**

[Returning\\_Data\\_to\\_the\\_Repository.pdf](#) (Adobe PDF)

#### 8.1.2.7. Other Investigator Tools

Other online tools for investigators include:

- description of the SWAN study design;
- the SWAN Study timeline;
- brief descriptions of completed and current Repository projects;
- schematic of the Repository application and review process; and
- FAQ page.

## 8.2. TOOLS FOR REVIEWERS

A second database, the Repository Admin Center, was built to automate and streamline the review of Repository applications. When a full application is submitted for SRO or SRAG review, individual reviewers can be selected online and a link to the application and review questions is automatically emailed to each reviewer. Reviewers can answer the objective review questions with a click of a button indicating his or her level of agreement, and provide comments on each point within the associated text boxes. Reviewers are asked to provide an overall recommendation and final comments before submitting online to Repository staff. When all reviews are submitted, the system automatically averages responses and compiles reviewers' comments and votes providing neat summaries for the applicants. (For more information on how this website and Application Reviews fit into the full Application and Review process, see section 6 of this manual.)

### Admin Center screenshot: Reviewer Questions

**Application Review: “Example Inquiry/Application Status”**

**PAGE 1**

There are no documents associated with this inquiry.

**QUALITY OF THE SCIENCE**

THERE ARE CLEARLY DEFINED HYPOTHESES OR RESEARCH QUESTIONS WITH SPECIFIED EXPOSURE AND OUTCOME INFORMATION.

01=Strongly Agree  
 02=Agree  
 03=Neutral  
 04=Disagree  
 05=Strongly Disagree  
 00=Not Applicable

**Comment** (Limit 1000 characters, or approximately 150 words.)

THIS REPRESENTS NEW SCIENCE, NOT A CONFIRMATION OF EXISTING INFORMATION.

01=Strongly Agree  
 02=Agree  
 03=Neutral  
 04=Disagree  
 05=Strongly Disagree

### 8.3. INTERNAL TOOLS FOR REPOSITORY STAFF

#### 8.3.1. Application Management

On the same Admin Center system that handles the automated review process, the Application Management system is available to internal Repository staff. The *Application Management* site is a password-protected tracking system for all inquiries and requests to access SWAN Repository samples and data. It provides structure and space to store all communications and documents associated with each step of the life of an application – from the inquiry, application and review, resource transfer and data return, through the completion of the project. (<http://applicationmanagement.swanrepository.com>)

#### 8.3.2. SAS Code Generation from the Data Warehouse

The Data Warehouse also contains useful tools for Repository staff. The DW is built using a MySQL database, and links to the SAS (Statistical Analysis System) platform. Because frozen SWAN datasets are stored on the server as SAS datasets, this link makes possible the automation of SAS code generation, thus forging individualized datasets and codebooks. Investigators select SWAN variables of interest on the DW and save them into personal libraries. These variable lists drive the production of SAS code which pulls the selected data into datasets. This automation saves all users a great deal of time and resources.