

EXAMPLE NIH DATA MANAGEMENT PLAN 2023

DATA TYPE

Types and amount of scientific data expected to be generated in the project:

The digital data generated during the project are Excel spreadsheets with capacitance values that were measured using the Corneometer® 825 probe and TEWL values that were measured using the Tewameter® TM300 probe, coupled to the Courage & Khasaka MPA-5 equipment. Data will be collected from 1,000 research participants, generating 5,000 datasets totaling approximately 100 Gigabytes (Gb) in size. The following data files will be used or produced in the course of the project: Excel data files for initial collection, and transformed into comma separated value (.csv) files for optimal reuse in data repository. Raw data will be transformed by SPSS software and the subsequent processed dataset used for statistical analysis. To protect research participant identities, aggregated data will be made available for sharing.

Scientific data that will be preserved and shared, and the rationale for doing so:

The data will be stored in Research Data Repository of Unicamp (REDU). Data from 100 subjects with healthy skin, that were collected in a clinical study carried out to evaluate the effect of humectants used in cosmetic products in improving skin hydration and skin barrier efficiency, from capacitance measurements obtained with the Corneometer probe ® 825 and TEWL measurements obtained with the Tewameter® TM300 probe, respectively. These measurements will be shared in REDU after processing and de-identifying the raw data. Data that can identify research participants included in the study will be anonymized. Raw data may only be shared with the authorization of the Responsible Investigator, upon request made by the interested party.

A brief listing of the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate interpretation of the scientific data.

Data will be stored in XLS/XLSX format (spreadsheets) and can be accessed and manipulated using Microsoft Excel. When publishing in a repository, we will transfer the Excel spreadsheets into the open-source comma-separated value .csv format. To facilitate interpretation of the data, metadata, documentation, protocols, and data collection instruments will be shared and associated with the relevant datasets.

RELATED TOOLS, SOFTWARE AND/OR CODE

State whether specialized tools, software, and/or code are needed to access or manipulate shared scientific data, and if so, provide the name(s) of the needed tool(s) and software and specify how they can be accessed

Tabular formatted data will be made available in .csv format and will not require the use of specialized tools to be accessed or manipulated.

If applicable, specify how needed tools can be accessed, (e.g., open source and freely available, generally available for a fee in the marketplace, available only from the research team) and, if known, whether such tools are likely to remain available for as long as the scientific data remain available.

The Excel tool, which can be used to create datasets and spreadsheets, is available free of charge through my university. SPSS will also be used. SPSS is a proprietary software product owned by IBM. To facilitate the re-use of this data, one may need to purchase a license for access to SPSS to do statistical analysis.

STANDARDS

State what common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources, and provide the name(s) of the data standards that will be applied and describe how these data standards will be applied to the scientific data generated by the research proposed in this project. If applicable, indicate that no consensus standards exist

Whenever possible, we will use common data elements to structure and organize our data. Our numerical data will be structured and described using the Darwin standard, which has been widely adopted in the dermatological community.

DATA PRESERVATION, ACCESS, AND ASSOCIATED TIMELINES

Repository where scientific data and metadata will be archived:

Dataset(s) resulting from this research will be shared via the generalist repository Dryad, which provides metadata, persistent identifiers (i.e., DOIs), and long-term access. Dryad is the institutional data repository supported by the University of California and all data is shared under a CC0 waiver, which makes the dataset(s) publicly available.

Data will be made available as soon as possible or at the time of associated publication. Dryad datasets are backed up to Merritt, the UC's CoreTrustSeal-certified digital repository, for long-term storage and accessibility. Procedures in place to ensure dataset preservation include storage of data files in multiple geographic locations, regular audits for fixity and authenticity, and succession plans in the event of repository closure.

How scientific data will be findable and identifiable:

Dryad provides metadata, persistent identifiers in the form of DOIs, and long-term access. Dryad is the institutional data repository supported by the University of California and all data is shared under a CC0 waiver, which makes the dataset(s) publicly available. Data will be made available as soon as possible or at the time of associated publication. Dryad datasets are backed up to Merritt, the UC's CoreTrustSeal-certified digital repository, for long-term storage and accessibility.

When and how long the scientific data will be made available:

Data will be made available as soon as possible or at the time of associated publication.

ACCESS, DISTRIBUTION, OR REUSE CONSIDERATIONS

Factors affecting subsequent access, distribution, or reuse of scientific data: NIH expects that in drafting Plans, researchers maximize the appropriate sharing of scientific data. Describe and justify any applicable factors or data use limitations affecting subsequent access, distribution, or reuse of scientific data related to informed consent, privacy and confidentiality protections, and any other considerations that may limit the extent of data sharing. See [Frequently Asked Questions](#) for examples of justifiable reasons for limiting sharing of data.

There are no limiting factors affecting subsequent access, distribution, or reuse of scientific data.

Whether access to scientific data will be controlled: State whether access to the scientific data will be controlled (i.e., made available by a data repository only after approval).

No scientific data will be controlled.

Protections for privacy, rights, and confidentiality of human research participants:

If generating scientific data derived from humans, describe how the privacy, rights, and confidentiality of human research participants will be protected (e.g., through de-identification, Certificates of Confidentiality, and other protective measures).

The study protocol, under the CAAE number of 29880820.2.0000.5404, was approved by the Ethics in Research Committee (CEP) according to report number 4.018.746. The study was conducted in accordance with the Good Clinical Practice and the Declaration of Helsinki. All subjects provided written informed consent. To protect participant privacy and confidentiality, shared data will be de-identified using the masking method.

OVERSIGHT OF DATA MANAGEMENT AND SHARING

Describe how compliance with this Plan will be monitored and managed, frequency of oversight, and by whom at your institution (e.g., titles, roles).

The following individuals will be responsible for data collection, management, storage, retention, and dissemination of project data, including updating and revising the Data Management and Sharing Plan when necessary.
Anna Laboratory, Data Manager, UNMC, ORCID, alab@unmc.edu

PLANNED RESEARCH OUTPUTS

Dataset for humectants

PLANNED RESEARCH OUTPUT DETAILS

Title	Type	Anticipated release date	Initial access level	Intended repository	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
humectants	Dataset	2028-01-24	Open	DRYAD	50 MB	CC Attribution Non Commercial Share Alike 4.0 International	Dublin Core	Yes	No