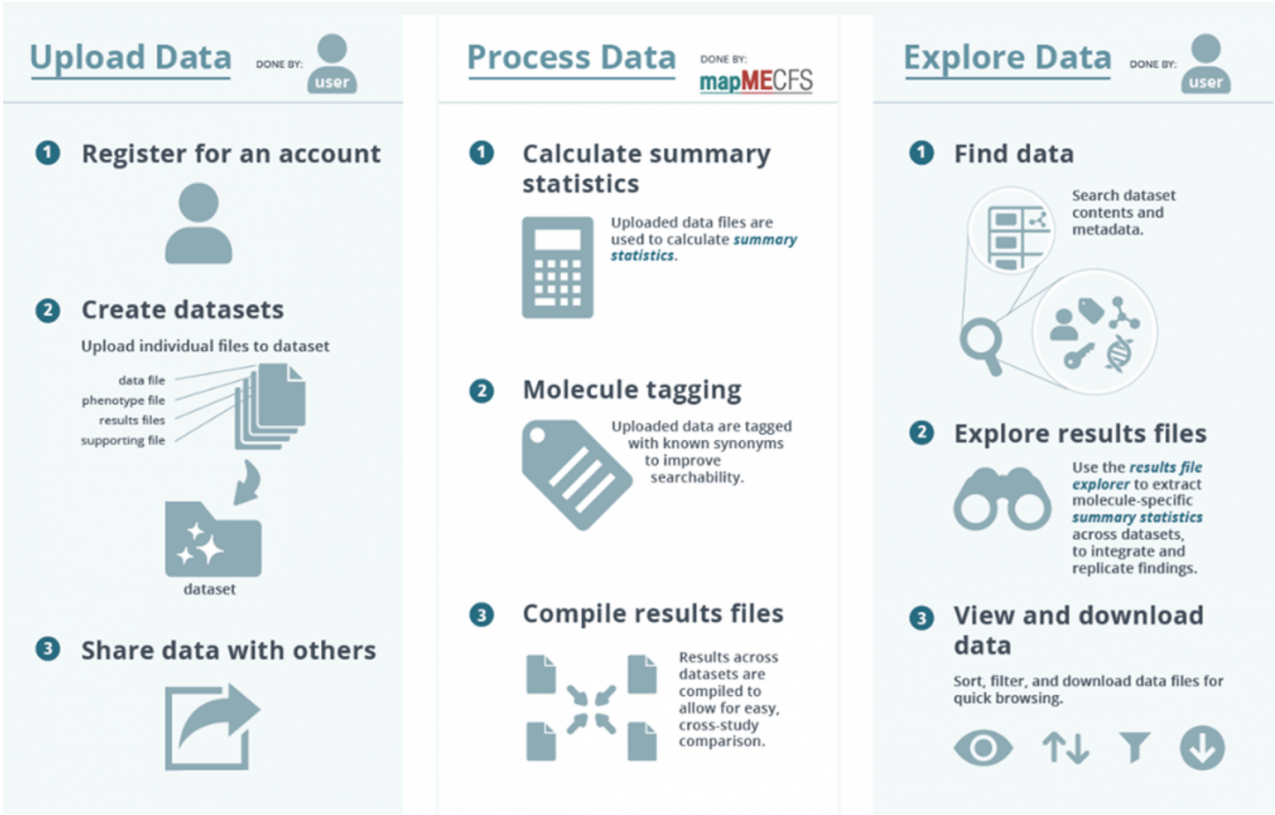
Data and specimen-sharing tools offer new discovery opportunities for ME/CFS researchers

openaccessgovernment.org/article/data-and-specimen-sharing-tools-offer-new-discovery-opportunities-for-me-cfs-researchers/174003/



Within the field of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) research, two online tools: mapMECFS (Mathur and Carnes, 2021) and searchMECFS play a crucial role in advancing the understanding of ME/CFS by encouraging researchers to share and use data and biospecimens that are stored in centralized and easily accessible data portals

Overview

mapMECFS and searchMECFS are hosted by RTI International and funded by the National Institutes of Health (NIH). mapMECFS is the largest interactive portal and repository for ME/CFS data. It offers researchers from diverse disciplines a platform to share and integrate data from their ME/CFS research studies. searchMECFS is an interactive search tool that allows researchers to identify and request available biospecimens to conduct novel experiments.

Purpose and significance

mapMECFS aims to overcome the challenge of fragmented data sources in ME/CFS research by providing access to research results across many scientific disciplines and body systems. mapMECFS offers new opportunities for researchers by providing a centralized repository and tools to connect databases and enable exploration of complex study results. mapMECFS also serves as the repository for all experimental data generated from the use of biospecimens accessed through searchMECFS.

searchMECFS addresses the logistical challenge of biospecimen selection and access by assisting with the selection of biospecimens based on demographic and clinical attributes. Currently, searchMECFS houses demographic and clinical data from the Chronic Fatigue Initiative study and information about associated biospecimens stored in a central biorepository. Future plans include adding biospecimens from the Center for Disease Control and Prevention's Multi-Site Clinical Assessment of ME/CFS study.

User registration and access

mapMECFS user registration is open to any researcher planning to use the data for research purposes. The secure registration process involves agreeing to the Data Use Agreement (DUA) and submitting a registration form that is reviewed by Program staff at NIH. Approved new users will either be added to an existing mapMECFS Organization or to a user-specific Organization created for them. Researchers approved to upload data will also be asked to complete a User Agreement.

To ensure secure data access, mapMECFS incorporates a two-tiered dataset structure. Private datasets, characterized by restricted visibility, are only accessible to users within the same Organization. Private datasets allow researchers to upload and prepare their data while they await the publication of the supporting manuscript. In contrast, public datasets are accessible to all approved users, fostering a collaborative and inclusive environment for research exploration.

Once a dataset's supporting manuscript has been published, users are highly encouraged to make their results publicly available to registered users.

The registration process for searchMECFS mirrors that of mapMECFS, requiring similar information. Following approval, users can query the available biospecimens and associated demographic and clinical data. Once users identify their desired samples, they follow the link to the Biospecimen Resource Access Committee Application webpage where they find instructions and links to request the biospecimens of interest.

Data types and formats

mapMECFS supports a wide array of data types including proteomics, metabolomics, methylation, gene expression, microbiome, demographic, and health and survey data. The platform also accommodates supporting phenotype (clinical and demographic data) and data dictionaries. Data must be deidentified, so that all participant information is protected.

While formatting requirements vary for each data type, detailed documentation on the website guides researchers through the submission process, ensuring data consistency and integrity. The mapMECFS support team is also available to assist researchers with their data submissions.

Search and discovery

mapMECFS offers several features designed to assist researchers in accessing relevant data. These include (1) a user-friendly data explorer tool that enables researchers to effortlessly search for datasets containing specific analytes of interest and (2) a data integration tool which allows researchers to seamlessly merge clinical data files from specific cohorts.

searchMECFS provides an interactive query tool that can be used to build and execute data queries to find available biospecimens that meet the user's specified search criteria. Users can add multiple search criteria to further enhance their search and refine their results.

Privacy and data standards

Participant privacy, a main priority, is covered within the DUAs and User Agreements. All data uploaded on mapMECFS must be free of Personally Identifiable Information (PII). Agreements also cover responsible data use including refraining from attempting to identify participants, safeguards against unintentional disclosure, prompt reporting of any unauthorized use, and restrictions on using the data for clinical or medical purposes. It is the uploader's responsibility to ensure that PII is not present, participant privacy is fully protected, and sharing is compliant with all other governing policies (e.g., IRB-approved protocols, embargos).

Future directions and user engagement

The development of mapMECFS is guided by user input, prioritizing site enhancements according to the feedback received from researchers actively engaging with the platform. All site users are encouraged to contact the mapMECFS support team with suggestions to improve the site. RTI plans to expand the clinical and biological data within the site's integration tool and improve data standards and automated quality assurance pipelines. RTI is also working with the ME/CFS research community to expand the number of datasets publicly available in mapMECFS and to increase the number of biospecimens available through searchMECFS.

Funders: US National Institutes of Health (NIH): National Institute of Neurological Disorders and Stroke (NINDS); National Institute of Allergy and Infectious Diseases (NIAID); National Heart, Lung, and Blood Institute (NHLBI); National Center for Complementary and Integrative Health (NCCIH); and National Institute on Alcohol Abuse and Alcoholism (NIAAA)

Acknowledgement: This work was supported by the National Institutes of Health (NIH) under award number U24NS105535. The contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH or RTI International.