

**Center for Neuroscience and Regenerative Medicine (CNRM) Data Repository:  
Contribution and Access Policies  
4-16-12**

The CNRM Data Repository will contain data of human subjects (patients and non-patient populations) from observational through interventional clinical studies. The policies in this document are tailored to human subjects research data and are not set up for animal studies at this time. The personnel of the CNRM Informatics Core will develop and maintain functions within the CNRM Data Repository. An appendix to this policy document includes an explanation of these entities along with definitions of the terms and abbreviations used within this document.

**(1) Centralized Database.** Data from all CNRM-supported human subjects research should be deposited in the CNRM Data Repository, maintained by the Informatics Core. Investigators will be informed of this requirement to share data prior to accepting funding from the CNRM and the information will be provided with funding announcements when possible. CNRM studies initiated prior to establishment of the CNRM Data Repository are strongly encouraged to follow these policies as well.

**(2) IRB and Consent.** All CNRM-supported human subjects research should include in the Institutional Review Board (IRB) protocol and consent form that de-identified and coded data using Global Unique Identifier (GUID) will be deposited in the CNRM Data Repository and will be available as requested through the CNRM Data Quality, Access, and Publication Committee (see 5). Sample language for inclusion in IRB submission will be provided by CNRM program managers. The key allowing re-connection of the data to identifying information will reside only with the protocol Principal Investigator (PI). Permission should be requested to submit data to the CNRM Data Repository and associated Federal Interagency Traumatic Brain Injury Research Database (FITBIR). The PI should comply with any local IRB requirements for review of requests by investigators to access the data through the CNRM Data Quality, Access, and Publication Committee.

**(3) Common Data Elements (CDE).** All CNRM-supported human subjects research should use the NINDS TBI CDE to the maximal extent possible. The CNRM Informatics Core will consult with the PI of each study during the design of the research database for each study to make sure that it is maximally CDE compliant. The Informatics Core will work with the PI to map variables collected through Case Report Forms used in each study into CDE compliant variables.

**(4) Data to be deposited in the CNRM Data Repository.** All data from CNRM supported studies should be deposited in the CNRM Data Repository. Data will include CDE variables, metadata and individual patient data from imaging, biomarker, or physiologic studies. In addition to depositing the research data, the PI should also deposit the Study Protocol, Manual of Operations, Case Report Forms, and Data Dictionary. This ancillary information is essential for maximal utility of the research data. Ultimately, this resource will be helpful to the investigator as a user-friendly platform that is compliant with IRB and ethical considerations for performing analyses.

NOTE: For studies that are supported in part by **non-CNRM** sources, the funding source may have restrictions that do not permit compliance with the policies of the CNRM Data Repository or FITBIR to accomplish full data sharing among investigators. In such cases, the Data Quality, Access, and Publication Committee (see 5) will determine the applicable rules and policies in collaboration with the outside funding agencies on a case by case basis, with deference to pre-existing policies, regulations, and constraints.

**(5) Data Quality, Access, and Publication Committee.**

**(a) CNRM will create a Data Quality, Access, and Publication Committee.** This committee will review policies of the CNRM Data Repository to optimize the functionality and ensure compliance. The committee will review requests for data access. The committee will maintain a record of publications generated from data in the Repository and review plans for future publications as relevant to requests for data access.

**(b)** The committee members will be appointed by the CNRM Directors. The CNRM Informatics Core PI and CNRM Director of Clinical Research will co-chair the committee. Membership will include approximately five CNRM investigators that can represent PIs collecting data at WRNMMC, NIH, and other collaborating sites. Members may be added or replaced as needed. The committee will include representatives from NIH and USUHS, who will have experience with policies and procedures applicable to studies funded by intramural and extramural NIH and the Department of Defense as well as those who may be involved with funding from other sources that commonly share support of studies using CNRM resources. The committee will include representatives who have experience with clinical trials, epidemiology, and informatics. The committee will also include those with scientific or regulatory responsibility for the CNRM Data Repository.

**(6) Time of Data Deposition.** Data will be deposited in the CNRM Data Repository after it is cleaned and locked by the PI for each study. The PI is ultimately responsible for determining when the dataset has been completed and cleaned. Data quality must be sufficient to comply with Good Clinical Practices standards.

**(a) Time of data locking for ongoing studies with publication prior to end of study.** For ongoing studies, the time of data locking is when a paper is accepted by a journal based on that data. That locked version of the database then becomes available for others to request to replicate the analysis.

As these studies are ongoing, additional data can be added, including in some cases correction of errors discovered in earlier versions of the database. These corrected errors should be clearly marked in the ongoing active database, but once a database used in a given publication is locked, it should not be changed, even if errors are discovered. An audit trail of changes will be provided to assist interpretation of the changes for future analyses.

**(b) Time of data locking for completed studies.** For studies with a fixed performance period, the time of data locking is 6 months after the end of the study. For studies with a fixed participant enrollment target, the time of data locking is 6 months after the last participant has completed the last scheduled visit. If data locking is not possible within this 6 month period, extensions can be requested through the Data Quality, Access, and Publication Committee but will only be granted under special circumstances with clearly documented justifications.

**(7) Availability of data to other investigators.** The purpose of the CNRM Data Repository is to catalyze and facilitate research by making CNRM study data available to the larger community of qualified investigators. This is a priority of several funding agencies and is the best use of the scarce financial and participant resources used in these studies. At the same time, multiple national databases provide an embargo period to recognize that the PI who collected the data should have some degree of privileged access, to allow time to complete the primary analysis as well as important secondary analyses.

**(a)** Data in the CNRM Data Repository will have a one year embargo starting from the time the data is locked. While the data is available for query and download by qualified investigators from the moment it is locked, during the one year embargo only the PI for the project (or his/her collaborators/designees) can use that data for manuscript submissions. For special circumstances, e.g. PI deployment, extension of the embargo period can be requested through the Data Quality, Access, and Publication Committee but will only be granted under special circumstances with clearly documented justifications. Any request can only extend the embargo for up to one year so that any further extensions must be submitted as a new request for review.

**(b)** Data from the CNRM Data Repository will be open to access through the FITBIR system during the one year embargo period only to researchers who have submitted data into the FITBIR. At the end of the one year embargo period, the data will be open to all qualified researchers regardless of prior submission of data into the CNRM repository or FITBIR.

**Table 1: Data Access Periods**

Investigators	CNRM Data Management			CNRM Data Repository	
	Study Open	Study Ended	Data Lock < 6 months	Data Embargo 1 year	Full Data Access > 1 year
PI					
CNRM					
FITBIR Data submitters					
Other Qualified Researchers					

**(8) Requesting Data Access.**

**(a) Preview of Summary Demographic Data.** Prior to obtaining full IRB approval and formally requesting data access, investigators may wish to have a preliminary way to evaluate the appropriateness of the available dataset. Queries will be supported for summary non-identifiable demographic information, diagnosis, and descriptions of available data. The Informatics Core will provide this summary demographic data on a secured website and will register and verify investigators before allowing query access. Investigators will not be required to have IRB approval or permission from the Data Quality, Access, and Publications Committee. The ability to carry out such queries by the broad scientific community will allow investigators to evaluate whether the available data is appropriate for their purposes and therefore facilitate the use of the data and avoid processing unnecessary requests for data access. This policy is similar

to those used by the NIH intramural BTRIS database, as well as the NDAR, NACC and dbGAP databases.

**(b) Requests for access to CNRM Data Repository.** The Data Quality, Access, and Publication Committee will be responsible for procedures for investigators to provide the necessary information for review of investigator qualifications and data usage. The committee can revise the request procedures as needed to improve function or comply with regulatory requirements. Investigators requesting access to the CNRM Data Repository will prepare a brief (1 paragraph) description of the research question under investigation. The request will also include a list of all investigators and collaborators who will have access to the data with documentation that they have been certified in the Ethical Conduct of Research and Human Subjects Protection Training. Investigators must submit documentation of IRB approval of the research project with consideration of approvals across multiple sites if applicable.

Only de-identified data can be requested. The PI for each study is the only keeper of the code allowing identifying information to be connected to the data. The PI is responsible for all aspects of the code, including duration for which the code will be maintained.

Use of the data is limited to the project that was proposed and approved. The data may not be reused for other projects or analyses, or redistributed to other investigators, without approval from the CNRM Data Quality, Access, and Publication Committee.

**(c) Mechanism for review of requests to access data.** The Data Quality, Access, and Publication Committee will review requests within 2 weeks of receipt of full documentation. The committee will ensure that the requesting investigator has the proper IRB and regulatory permissions. The IRB approvals and protocol should provide evidence of scientific review of the goals of the proposed study in which the data will be used.

Requests will also be reviewed for redundancy and overlap with previously approved requests. However, redundancy or overlap with previously approved requests will not be automatic reasons for denial. Rather, if the request is to be approved, the committee will inform the requesting investigators that a data request for a similar project was granted (or was in the process of being granted). In addition, the PI who deposited the data will be provided an opportunity to respond to the committee, if applicable, to demonstrate progress toward similar goals. The purpose of this review is to minimize duplication of effort.

**(d) Notification of protocol PI when his or her data has been downloaded.** The PI of each protocol will be notified when data from his or her study was downloaded. This notification will include the name of the downloading investigator, contact information, and title of the data request. The notification will not include the specific hypotheses being tested or other detailed descriptors of the project.

**(e) PI responsibilities following data access.**

Investigators using CNRM data in publications should acknowledge CNRM and other relevant funding sources in their publications. Researchers who use CNRM data in their manuscripts must submit their work to Data Quality, Access, and Publication Committee before submission for publication. This is required to ensure that CNRM-related manuscripts are tracked and that CNRM funding is credited. Manuscripts will not be primarily reviewed for scientific content. Authors will receive feedback within 2 weeks of submission. The authors must submit a final accepted version of each manuscript to the Data Quality, Access, and Publication Committee. It remains the author's responsibility to work with the publisher to fulfill obligations of agency funding the work, such as obtaining a [PubMed Central ID](#), in compliance with NIH policy on public access.

Investigators accessing de-identified data are prohibited from any form of re-identification of data.

**(f) Compliance.** The Data Quality, Access, and Publication Committee will be responsible for overseeing compliance with policies of the CNRM Data Repository. Investigators who are identified as having failed to adhere to the designated policies will be blocked from further access to the Data Repository. Infractions considered by the committee as requiring more serious actions can include recommendation to the source of funding to discontinue support to the investigator or for the relevant study.

**(9) Coordination with FITBIR.** The CNRM Data Repository is developed to align effectively with the FITBIR. FITBIR has extensive documentation of policies and procedures currently under review. For issues that are not specifically addressed in this CNRM document, the FITBIR documents will serve as the default standards for use of the CNRM Data Repository. Once final FITBIR documents are available, the CNRM policy will be altered if necessary to comply with the FITBIR data sharing and access policy.

**Appendix: Terms and Definitions:**

Term	Definition
CNRM Data Repository	An informatics system with collective of de-identified clinical related research data resulting from studies funded within the Center for Neuroscience and Regenerative Medicine (CNRM).
Global Unique Identifier (GUID)	GUID is a unique computer generated alphanumeric code using a set of pre-defined personal identifiers entered only by study investigators and is used solely by the data repository to associate same subject data across different studies. GUID is a one-way hash code that cannot be reserved to re-identify the study participant. The identification keys is held only by the study PI. GUID indeed is a “de-identified and coded” identifier.
Cleaned Dataset	The clinical research dataset that were re-checked to ensure the data quality before submitting to data repository; cleaned data sets could be changed in the data repository if mistakes are found. Changes will be marked with audit trails.
Locked Dataset	The clinical research data sets that were submitted and accepted for publication or the trial data sets that were locked by data monitors, e.g. CRO. Locked datasets should never be changed in the data repository.
Informatics Core	The informatics core is part of CNRM scientific cores that provides essential informatics services such as electronic data capture and reporting for clinical protocols, participation in national TBI research and data repository community, research computing infrastructure, integration of CNRM technology requirements, and maintenance of CNRM central data

	repository.
Common Data Element (CDE)	A scientifically-vetted and standardized set of data <i>elements</i> and report forms using content standards that enable clinical investigators to systematically collect, analyze, and share data across the research community.
Data Quality, Access, and Publication Committee (DQAPC)	CNRM DQAPC will review CNRM Data Repository operational policies, data access requests and publications involving the data.
Federal Interagency TBI Research Database (FITBIR)	The FITBIR database is an informatics system and central data repository developed by DOD and NIH to store and link together de-identified phenotypic, diagnostic, treatment and outcome data derived from individuals who participated in TBI research studies.
Good Clinical Practice (GCP)	GCP is a standard for the design, conduct, performance, monitoring, auditing, recording, analyses, and reporting of <b>clinical trials</b> that provides assurance that the data and reported results are credible and accurate, and that the rights, integrity, and confidentiality of trial subjects are protected. Ethical and scientific quality standards for designing, conducting, recording and reporting trials that involve participation of human subjects to ensure that the RIGHTS, SAFETY and WELL BEING of the trial subjects are protected. Ensure the CREDIBILITY of <b>clinical trial</b> data. <a href="http://ichgcp.net/">http://ichgcp.net/</a>
Principle Investigator (PI)	PI is the individual(s) judged by the applicant organization to have the appropriate level of authority and responsibility to direct the project or program supported by the grant and who is (are) responsible for the scientific and technical direction of the project. (new NIH definition, see <a href="http://www.gpo.gov/fdsys/pkg/FR-2009-11-10/html/E9-27025.htm">http://www.gpo.gov/fdsys/pkg/FR-2009-11-10/html/E9-27025.htm</a> )