include

International Network for Cross-Linguistic Research on Brain Health

Data Management Guidelines

Include Network Data Management Guidelines

1. Rationale

This document outlines **Include**'s guidelines on data management. **Include** will advance novel cross-linguistic/cross-dialectal protocols harmonized across the languages of its participating institutions. The goal is to glean original, high-quality data to identify empirical patterns that are common across different languages/dialects as well those that are language- or dialect-specific.

Data is hereby defined as all information pertaining to, but not limited to, the following: speech and/or language samples and/or outcomes as well as demographic, clinical, neuropsychological, neuroimaging, and/or biofluid measures. This includes the raw data and data derived from their analyses.

2. Concept Sheets

2.1. Submitting a Concept Sheet

Include members will be invited to present Concept Sheets, that is, standardized documents required to formally propose a study (leading to publications) based on the network's collective data. Our official template can be requested via e-mail: include-network@gbhi.org.

Concept Sheets involve an explicit agreement among all parties involved in the project concerning roles, authorship, timelines, and other relevant considerations, always in accordance with these Data Management Guidelines as well as the Publication Guidelines. The standardized Concept Sheet asks prospective authors to specify the principal hypotheses, variables considered, and analysis plan. In order to advance a Concept Sheet, members should have contributed in some way to the Include data, as outlined in the Membership Guidelines. Options include contributions of subject data/datasets, expertise in task harmonization and validation, and/or expertise in relevant data analysis techniques. These contributions will be evaluated to ensure equitable access to the data. Members should then access and submit the Concept Sheet to the Coordinating Committee, at include-network@gbhi.org. The completed Concept Sheet will then be routed to the Coordinating Committee and evaluated based on the criteria in the next sections.

2.2. Evaluation of Concept Sheets

Concept Sheets will be reviewed and assessed by the Coordinating Committee based on the following criteria:

- scientific merit and feasibility;
- appropriateness of the investigator's qualifications and resources to protect the data;
- appropriateness to the goals of the project and those of **Include** at large;
- potential overlap with ongoing analyses; and
- compliance with previous agreements in the case of researchers who have previously accessed data from **Include**.

Based on these criteria, the Coordinating Committee will reach one of the following decisions: (i) accept proposal, (ii) request amendments for re-evaluation, or (iii) reject proposal. **Include** will also help members with approved Concept Sheets to find theoretical, methodological or technical support across the network to accomplish proposed studies.

2.2.1. Accepted Concept Sheets. Accepted Concept Sheets will lead to specific communication between the requesting member(s) and the Data Governance Committee, who will facilitate data access and indicate allowed and disallowed uses. After a Concept Sheet is approved, the investigators involved must sign a Data Management Agreement that is specific to the project. Thereupon, the de-identified data will be made available to investigators to conduct analyses. All analyses will be based on datasets that have been prepared, cleaned, and frozen periodically.

2.2.2. Concept Sheets requiring amendments. These Concept Sheets will re-enter the evaluation cycle once they have been revised and formally resubmitted to the Coordinating Committee. Concept Sheets resubmitted 6 months after the previous submission time will be reviewed as new submissions.

2.2.3. Rejected Concept Sheets. Rejected Concept Sheets will not be eligible for resubmission. Participating members, however, will be encouraged to delineate new Concept Sheets and submit them separately.

2.3. Securing IRB approval

Each member will need to secure IRB approval in compliance with international and local laws, including explicit authorization for open data sharing with third parties pending proper permission requests. Sites will only be fully enrolled once they have secured proper IRB approval.

3. Consideration for projects based on newly collected (not pre-existing) data

3.1. Accessing data collection tools

Projects may require data collections tools that are not available at specific participating sites. In these cases, members will partake in an **Include**-hosted video call to jointly decide on how to provide access to such tools, considering potential costs and logistic demands involved. Should specific tools be shared with or shipped to a given center, an explicit agreement will be signed stipulating the cost-bearer(s), allowed and disallowed uses, and return of materials, if applicable.

3.2. Data collection

Each **Include** project will rest on an explicit data collection protocol to be followed by all enrolled sites. If needed, project-leading sites should offer training and/or clarification alternatives for enrolled sites, aiming to maximize harmonization, efficiency, and data quality.

The Concept Sheet will specify stipulations concerning the type and amount of data to be collected at each site, as well as the timeline for its collection and considerations regarding its management, de-identification, and storage. Each project will indicate whether individual sites are at liberty to use the data they collected for purposes other than those pertaining to the cross-linguistic or cross-dialectal project proper.

4. Data sharing and storage

4.1. Basic considerations

Data sharing and storage procedures will be established jointly by all parties involved in a new project, during a meeting coordinated by **Include**'s Coordinating Committee.

These procedures will follow the policies mandated by the IRBs and regulations of the collaborating institutions. Data sharing and storage procedures should be performed in compliance with privacy regulations, human subject protection guidelines, material transfer agreements, and state and federal laws.

4.2. Data confidentiality

All precautions to ensure confidentiality must be taken by recipients of **Include** data. Datasets will be stripped of identifiers prior to release for sharing and transferred only with encryption and password protection by the **Include** Data Management Team. The code linking a subject's identity to data will be maintained in a secure place and will only be accessible to research staff on a need-to-know basis.

4.3. Adherence to local and national policies

All data sharing will follow the national policies of local sites. We will work with each site to ensure that all research participants have consented with allowance for open data sharing for research use. Individuals who are unable to provide consent for data sharing will not be included in shared datasets.

5. Data use

5.1. Fair use

Requested datasets may not be used for purposes other than the ones outlined in the Concept Sheets. Explicit permissions should be obtained from the Coordinating Committee for unanticipated uses.

5.2. Reciprocal support

Results generated through analyses of **Include** datasets must be returned to the Coordinating Committee for possible inclusion in the project database. A six-month embargo will be placed on returned data to allow publication of results. Notification of negative findings should be returned to the Coordinating Committee within one year of data set distribution.

5.3. Sharing of data with third parties

No sharing of **Include** data with a third party is allowed without permission of the Coordinating Committee.

6. Requesting expert support

Sites enrolled in an **Include** project may request expert support from other network members in order to achieve the aims of their approved Concept Sheets. **Include** will facilitate contact with relevant specialists to meet theoretical, methodological, and/or technical needs (e.g., for analysis or manuscript preparation purposes). Requests should be submitted to the Research Committee via e-mail, with a detailed account of needs, timelines, and benefits for interested parties.

7. Additional considerations

7.1. Authorship and acknowledgments

Please refer to **Include's** Publication Guidelines for recommendations on authorship and acknowledgments.

7.2. Communication of publications to Include

Acceptance of manuscripts and presentations (as defined in the Publications Guidelines) that have utilized **Include's** shared data should be communicated to the Dissemination and Outreach Committee for internal documentation and dissemination. **Include** retains the right to disseminate papers and presentations based on **Include** data.