

BD² Open Science Policy

Version 4: July 2026

Summary

Open science is core to the mission of BD² and a critical component of accelerating the progress toward our vision that all people with bipolar disorder thrive. All research supported by BD² must adhere to the BD² Open Science Policy, which was developed using guidelines set forth by several leading organizations and efforts (see Appendix B for more information and endorsements).

This policy applies to all funded investigators within the BD² network. Failure to meet these requirements may result in delayed funding or inability to receive extended funding.

Index

I. Preprint Requirement – Page 2

II. Manuscript Sharing Requirement – Page 3-4

III. Acknowledgement Requirement – Page 5

IV. Method and Code Sharing Requirement – Page 6-7

V. Data and Resource Sharing Requirement – Page 8-9

VI. Network Sharing Requirement – Page 10

Appendix A: Open Science Checklist – Page 11

Appendix B: Endorsements and Glossary – Page 12-13

Appendix C: Best Practices in Resource Sharing – Page 14-15

I. Preprint Requirement: To accelerate the dissemination of new scientific findings to the research community, all BD² Manuscripts¹ must be preceded by a preprint that is deposited to an open-access repository no later than the time the article is first submitted to a journal for consideration.

Preprint Server Selection: Preprints must be posted to an open-access repository that provides a persistent identifier, e.g., a digital object identifier (DOI). Acceptable preprint servers include bioRxiv and medRxiv; if a preprint is posted to bioRxiv or medRxiv, BD² may index it within our [bioRxiv/medRxiv Channel](#).

License Selection: Investigators must select a CC BY (or CC BY 4.0) license or a CC0 public domain dedication. “ND”, “NC”, and “SA” modifiers on a CC BY license are not permitted, as they restrict reuse (see Glossary in Appendix B for expanded definitions of license types).

Compliance Check: Once the preprint is posted, the persistent identifier and/or weblink associated with the preprint must be emailed to discoverygrants@bipolardiscoveries.org, which will trigger an Open Science Check. During this check, BD² will confirm that the correct license type was selected and provide a checklist of all components of the manuscript that are pursuant to the Open Science Policy. All items on the checklist must be made compliant prior to final formal publication of the manuscript, though investigators are encouraged to make the preprint itself as compliant as possible. To this end, investigators may request that BD² performs an Open Science Check on a draft prior to preprint deposition.

The preprint will additionally be reviewed after a BD² Manuscript is published in a peer reviewed journal. This secondary check is used to confirm that the preprint exists and that it was posted no later than the time of journal submission. These dates are available within the article metadata.

Minimum Requirements:

- All BD² Manuscripts must be preceded by an open-access preprint that is posted no later than submission of the paper to a journal.
- All preprints must have a persistent identifier (i.e. a DOI).
- All preprints must use a CC BY 4.0, CC0, or equivalent license type.
- All preprints must be sent to BD² for an Open Science Check.

Additional Recommendations:

- All preprints should be as compliant with the BD² Open Science Policy as possible, meaning that all datasets, protocols, and code are recommended to be publicly shared at the time of the preprint.
- Investigators may request an Open Science Check from BD² on a preprint draft prior to deposition.

¹ A BD² Manuscript is defined as any manuscript that results from an effort that was fully or partially funded by BD², including any manuscript that includes primary BD² Data or the use of BD² Biospecimens. See the Glossary in Appendix B for more details.

II. Manuscript Sharing Requirement: To promote broader sharing of scientific findings and foster transparency, all BD² Manuscripts must be fully open-access and clearly identify all resources generated and used.

Journal Selection. BD² Manuscripts must be published in a peer-reviewed journal and made immediately open access with no embargo period. Authors may choose either fully open access journals, such as eNeuro or BMC Psychiatry, or hybrid open access journals that allow authors to pay an Article Processing Charge (APC) to upgrade their manuscript to open access. For BD² funded investigators, BD² will cover the difference in fees to make a BD² Manuscript open access if and only if all terms of the Open Science Policy are met.

It is the responsibility of the authors of a BD² Manuscript to ensure their publication is open access. They may choose one of the following three routes:

1. Select a journal that allows for open access publication using the required license type (see below). The [Open Policy Finder](#) website can be used to find the relevant policies of most scientific journals.
2. Post the Author Accepted Manuscript (see Glossary in Appendix B) to a publicly accessible, community-recognized repository with no embargo period, e.g., PubMed Central or an institutional repository.
3. Post the Author Accepted Manuscript as a revised preprint.

License Selection. BD² Manuscripts must use a CC BY or CC BY 4.0 license, or CC0 public domain dedication (see Glossary in Appendix B for expanded definitions of license types).

If a journal does not have the appropriate licensing and/or is subscription-only, the submitting authors must select a different journal or a secondary publication venue, such as an institutional repository or revised preprint.

Key Resources Table. BD² Manuscripts must include a Key Resources Table (KRT), which provides details on all resources generated and/or used within a publication. Resources include all protocols, code, datasets, and laboratory tools (see Sections IV and V). The KRT must include descriptive labels and persistent identifiers for all resources; a template can be found [here](#). This resource index has two main benefits: (1) it makes resources easily findable and accessible to readers of the manuscript and (2) it allows for a quicker and more efficient Open Science Check.

After a KRT is generated, it must be included in the manuscript in at least one of two ways:

- (Preferred) As an upload to a public repository that provides a persistent identifier, which should itself be cited within the manuscript (e.g. in the Data Availability Statement). BD² recommends depositing to the BD² community space on [Zenodo](#). Note that uploading a KRT to a repository adds an additional benefit in that it can be versioned after it is published, therefore allowing links to be updated after a manuscript is published.
- As a table within the main text of the BD² Manuscript. A Star Methods table, as required by some journals, meets this requirement.

Compliance Check. To ensure Open Science compliance, BD² Manuscripts must be emailed to discoverygrants@bipolardiscoveries.org prior to *final* submission, i.e., before a paper is “fixed” and the journal will no longer accept edits. **If a BD² Manuscript has been published without an Open Science Check, authors risk being in violation of the Open Science Policy.**

At a minimum, authors must allow 7-10 business days for the Open Science Check to be completed and should plan for 2-3 weeks post check to incorporate any required updates. BD² recommends that the Open Science Check is initiated just before the point when authors expect an article to be accepted, i.e. after a manuscript has gone through reviews and revisions, but prior to receiving an Author Accepted Manuscript. Note that this will be at least the second Open Science Check, as a previous check will have been completed on the preprinted version (see Section I: Preprint Requirement).

During the Open Science Check, BD² will confirm that all terms of the Open Science Policy have been met (see Open Science Checklist in Appendix A). At the point of this “final” Open Science Check, all protocols, code, and datasets are expected to be deposited to repositories. Note that these resources may be hidden from public view until final

manuscript publication, and most repositories can incorporate an embargo period to this end. In this case, authors should provide BD² with a private “reviewer link” for the Open Science Check.

Any instances of non-adherence that are identified during the Open Science Check must be corrected prior to publication. If the authors believe that the results of the Open Science Check are incorrect, it is their responsibility to contact BD² and reconcile any differences. BD² will do an additional check after final publication to verify that all updates were made, as well as confirm that the article is open access and has the correct license type. Failure to pass the Open Science Check will preclude coverage of APC fees by BD² and may result in delays in continued funding or in the loss of eligibility for future funding opportunities.

Minimum Requirements:

- All BD² Manuscripts must be published in an open access format and be made immediately available with no embargo period.
- All BD² Manuscripts must use a CC BY 4.0, CC0, or equivalent license type.
- All BD² Manuscripts must include a KRT.
- All BD² Manuscripts must be sent to BD² for a “final” Open Science Check prior to final publication; all components of the BD² Manuscript that fall under the Open Science Policy must be fully compliant by the time of the “final” Open Science Check.

Additional Recommendations:

- Investigators may request additional and/or earlier Open Science Compliance Checks from BD² on draft versions of a manuscript. This can reduce the burden of making open science updates at the time of manuscript finalization.

III. Acknowledgement Requirement: To recognize BD² as a network and build associations across BD²-funded work, all preprints, BD² Manuscripts, and resource outputs must acknowledge BD² funding and affiliation.

Acknowledgement. All preprints, BD² Manuscripts, and research outputs must acknowledge BD² grant support. This acknowledgement should be included in the Acknowledgements section and BD² should be listed as a funder for platforms that include such an annotation field (e.g. bioRxiv, Zenodo, protocols.io). If there is no specified place for acknowledgments or funders, the notes or detail area of the record can be used.

BD² recommends the following acknowledgement language (items in bold must be customized):

This research was funded in [whole/part] by Breakthrough Discoveries for thriving with Bipolar Disorder (BD²; <https://ror.org/00z5dw933>) (Grant ID: XXXXXX).

For BD² Manuscripts that specifically use the BD² Integrated Network Data, authors must also use the following language:

Data used in the preparation of this article were obtained from the BD² Integrated Network study (Dataset Version: **INSERT VERSION #**) sponsored by Breakthrough Discoveries for thriving with Bipolar Disorder (BD²; <https://ror.org/00z5dw933>). To access the data, researchers must register for the BD² Data Platform at <https://bd2datarelease.indocpilot.io/>. The authors acknowledge the efforts of the BD² Integrated Network Consortium members. The full list of current BD² Integrated Network Consortium members is available here: <https://doi.org/10.5281/zenodo.21708193>. For up-to-date information on the BD² Integrated Network study, visit bipolardiscoveries.org.

This language can be adapted as needed for other datasets or resources, ensuring the inclusion of “Breakthrough Discoveries for thriving with Bipolar Disorder (<https://ror.org/00z5dw933>)” and key resource identifiers with version information.

Affiliation. To recognize that BD² is both a funder and a network, and to ensure that the public can identify members of the network, BD² must be included in author affiliations. Any current or former member of a BD²-funded team who is an author on a BD² Manuscript should have a BD² affiliation in addition to their home institution affiliation. Team members include anyone who is or has previously been an active participant in research efforts associated with a BD² project or efforts at an Integrated Network site.

The BD² affiliation should be listed as:

Breakthrough Discoveries for thriving with Bipolar Disorder (BD²), Easton, MD 21601

Author Identification. All BD²-affiliated authors must also include their Open Researcher and Contributor Identifier (ORCID) in preprints and BD² Manuscripts. This system provides a persistent identifier to an investigator and is used to connect each unique researcher to their publication record.

Minimum Requirements:

- All BD² Manuscripts and preprints must include BD² in the acknowledgements.
- All resources generated that are associated with BD² funding must include BD² in the acknowledgements.
- Any use of Integrated Network Data must use the specific language referenced above in acknowledgements.
- All authors who are current or former members of the BD² network must cite BD² affiliation and include their ORCID.

IV. Method and Code Sharing Requirement: To promote research transparency and reproducibility, the details of any experimental methods or analytical processes developed under BD² funding or used to collect or analyze BD² Data must be described in detail and shared publicly no later than the time of publication of an associated BD² Manuscript. In addition to transparency, all methods and code are considered to be resources and must be shared in a way that promotes reuse.

Protocols and Standard Operating Procedures (SOPs). Any experimental method used to collect or analyze BD² Data or BD² Biospecimens must be described in detail and made publicly available using a protocol sharing service, such as protocols.io or [bio-protocol](https://bio-protocol.org), or through a peer-reviewed methods paper. Each protocol must be open access, have a persistent identifier, and use the least restrictive license (CC BY 4.0, CC 0, or equivalent).

A protocol must be published for each method used in a BD² Manuscript; a description in a Methods section is not sufficient. The authors of each BD² Manuscript are responsible for making sure a published protocol is available for all methods, even if they are not responsible for the development of all methods.

To promote reproducibility and reuse, protocols should be detailed and written “recipe-style” with step-by-step instructions (see Appendix C for additional guidance). Further, each protocol should reflect the exact steps that were used to generate or analyze data in a given publication. In the case of a method that will be used across multiple publications, a single “parent” protocol may be created and published once but may need to be versioned or forked to reflect any differences across publications.

For researchers in the BD² Integrated Network, protocols related to Integrated Network Data must be emailed to integratednetwork@bipolardiscoveries.org to be reviewed centrally by the BD² team prior to publishing. Protocols must not include details that may compromise the privacy of study participants or impede study objectives. Investigators using Integrated Network Data must cite the Longitudinal Cohort Study Protocol (<https://pubmed.ncbi.nlm.nih.gov/41428166/>).

All protocols must be clearly cited and easily findable. To this end, authors must:

- Add each protocol to the Key Resources Table of a manuscript,
- Include citations for all protocols within the text of a manuscript – authors can cite each protocol separately or cite a single protocol collection that itself includes links to each protocol, and
- Post each protocol onto the Protocols Database on the BD² Hub (see Section VII).

Code. Any code used under BD² funding must be clearly identified. Within this policy, the term code is used broadly and applies to all scripts, macros, algorithms, software, etc., that are used to manipulate data, including data processing, data cleaning, data analysis, data visualization, and computation.

For existing code that was developed outside of BD² funding, either by the authors of the BD² Manuscript or by an external party, the resource must be cited with version information included, and, if available, a persistent identifier such as a DOI or a Research Resource Identifier (RRID).

Any code developed using BD² funding or trained using BD² Data must be made publicly available. An exception may be made for code that has protected IP. If researchers are unable to share a resource due to restrictions on IP or commercialization concerns, they must request a waiver *prior* to starting any related workflows within their BD²-funded work. An IP waiver may preclude access to BD² Data or BD² Biospecimens.

For newly developed code, sharing must be open source and must be deposited with a license that allows for distribution, modification, and unrestricted reuse; BD² recommends the MIT License (or another permissive license). Copyleft licenses are also permitted (e.g., GNU General Public License), and they are required if copying code published with a copyleft license. Code is expected to be well-annotated and include usage notes (see Appendix C for additional guidance).

When citing newly developed code in a publication, two versions of the code are recommended to be included:

- “Living” version - active version that may be further updated (e.g., GitHub URL; note this is *not* a persistent identifier).
- “Archive” version - point-in-time archive of the code that was used for the manuscript and has a persistent identifier (e.g., copy of code posted to Zenodo; see instructions [here](#)).

Authors are encouraged but not required to share existing, non-BD²-funded code in an open source format if they were responsible for developing the resource.

For researchers in the BD² Integrated Network, code that is used to analyze Integrated Network Data emailed to integratednetwork@bipolardiscoveries.org to be reviewed centrally by the BD² team prior to posting. Code must not include details that may compromise the privacy of study participants or impede study objectives.

All code must be clearly cited and easily findable. To this end, authors must:

- Add all code, including software, to the Key Resources Tables of a manuscript; and
- Post each newly generated set of code to the Code & Software Database on the BD² Hub (see Section VII).

It is additionally recommended that authors include a Code Availability Statement and citations for each set of code within the main text of their BD² Manuscript (see more details in Appendix C).

Minimum Requirements:

- All protocols used in relation to BD² funded work must be deposited in an open access repository and made publicly available no later than the time of publication of an associated BD² Manuscript.
- All newly developed code generated under BD² funding or through the use of BD² Data or BD² Biospecimens must be open source and made publicly available no later than the time of publication of an associated BD² Manuscript
- Persistent identifiers for all protocols and code must be generated and cited in any BD² Manuscripts for which they were used.
- For protocols and code related to Integrated Network Data, the BD² team must review all links *prior* to their public release.
- For each BD² Manuscript, all cited protocols and code must be listed in the Key Resources Table.

Additional Recommendations:

- Protocols and code that are listed in a BD² Manuscript but were not used in relation to BD² Data should be made publicly available and open access.
- Protocol and code sharing should follow best practices to promote re-use (see Appendix C).

V. Data and Resource Sharing Requirement: All resources developed under BD² funding must be made available to the research community for reuse and to promote rigor and transparency. Thus, all newly generated research outputs, including datasets and laboratory tools, must be shared in a publicly accessible repository no later than the time of publication of an associated BD² Manuscript.

Datasets. Datasets generated with the use of BD² funding must be made publicly available, unless they include sensitive human subject information. Each dataset within a BD² Manuscript must be open access, have a persistent identifier, and use a CC BY (or CC BY 4.0) license or a CC0 public domain dedication. Note that, for datasets, CC0 is recommended over CC BY to encourage inclusion of the data in larger analyses without citation confusion (see explanation [here](#)).

Non-human datasets must be shared to a public data repository that is recommended by the appropriate field for the associated data modality (e.g., DANDI Archive for neurophysiology data and Brain Image Library or BioImage Archive for microscopy images). These repositories must:

1. Enable immediate open access to data upon full publication of the manuscript,
2. Allow reuse with licensing (no more restrictive than CC BY 4.0 or CC0),
3. Assign a unique and persistent identifier to the dataset, such as a digital object identifier (DOI) or accession number that allows linking and citing,
4. Include metadata necessary to understand, assess, and replicate the reported study findings, and
5. Provide long-term preservation and storage of data that meets the [ISO's trustworthy digital repository standards](#).

Extra considerations must be given to human-derived datasets, including those generated using iPSCs and organoids. These datasets must be de-identified (such that they retain no direct or indirect identifiers and no protected health information (PHI)) and shared with a public data repository in a manner that is as open as possible but as closed as necessary. Sharing of human-derived datasets may involve additional levels of protection, such as the incorporation of a Data Use Agreement, and must adhere to the terms and conditions of all IRBs and participant agreements.

Human-derived datasets cannot be made publicly accessible if they:

- Could reasonably compromise the privacy of human participants,
- Could impede the objectives of a study, or
- Could compromise the privacy or confidentiality of researchers or collaborators.

Published datasets consisting of Integrated Network data will be de-identified and made publicly available via the Integrated Network's Pilot platform by formal request to BD².

For all open access datasets, sharing must meet the FAIR standard, meaning that they are Findable, Accessible, Interoperable, and Reusable. Datasets must also include clear labels and metadata and BD² will evaluate these criteria during the Open Science Check. The specific format and recommended repository for each dataset will vary based on the datatype and size; it is required for each team to create a custom Data Sharing Plan for their study. Additional guidance can be found in Appendix C.

All shared datasets must be clearly cited and easily findable. To this end, authors must:

- Add each dataset to the Key Resources Tables of a manuscript, and
- Post each dataset onto the Datasets Database on the BD² Hub (see Section VII).

It is additionally recommended that authors include a Data Availability Statement and dataset citations within the main text of their BD² Manuscript (see more details in Appendix C).

Research Tools. Any research tool used in the collection or analysis of BD² Data must be clearly identified with a persistent identifier such as a Research Resource Identifier (RRID), including existing research tools (research "inputs") and newly generated resources (research "outputs"). Resources with RRIDs include cell lines, mouse lines,

antibodies, biosamples, and some equipment and software; BD² recommends using the [Resource Identification Portal](#) for locating identifiers for research inputs cited in publications. It is also recommended to include supplier catalog numbers, though these are not considered persistent identifiers and may be insufficient in the identification of a key resource. If a research tool does not fall under the RRID system, BD² recommends using an available persistent identifier, such as Chemical Abstracts Service (CAS) numbers.

For newly developed research tools (research “outputs”) that are included in a BD² Manuscript, the originator of the resource must register the output with an appropriate registry to obtain a persistent identifier (e.g. RRID or DOI) where applicable.

For users of BD² Integrated Network Data, note that the ownership of any derivative product resulting from the manipulation, harmonization, aggregation, or analysis of that dataset will rest with BD².

All research tools must be clearly cited and easily findable. To this end, authors must:

- Include RRIDs for all research inputs and outputs within the text of a manuscript,
- Include RRIDs for all research inputs and outputs within the Key Resources Table of a manuscript, and
- Post any new research outputs onto the Lab Resources Database on the BD² Hub (see Section VII).

Minimum Requirements:

- Researchers must develop a personalized Data Sharing Plan at the outset of their BD² funded project.
- All BD² Data must be made deposited and made publicly available no later than the time of publication of an associated BD² Manuscript.
 - Human-derived datasets are an exception to this requirement, as sharing must take extra consideration of protecting participant identity and PHI.
- Data sharing must follow best practices and FAIR principles to promote re-use (see Appendix C).
- All lab tools generated under BD² funding must be registered on a public database to obtain a RRID.
- Persistent identifiers for all research inputs and outputs must be obtained and cited in any BD² Manuscripts in which they were used.
- All research inputs and outputs must be included in a Key Resources Table for all BD² publications.
- For research outputs related to Integrated Network Data, the BD² team must review all links *prior* to their public release.

Additional Recommendations

- Datasets that are included in a BD² Manuscript but were funded outside of the BD² Network should be made publicly available and open access.
- Lab tools that are used in a BD² Manuscript but generated outside of BD² funding should be registered in a relevant repository to obtain a RRID.

VI. Network Sharing Requirement: To promote sharing and organize BD² funded work, all research outputs generated fully or partially with BD² funding must be catalogued and shared within the BD² Network.

BD² Hub. The virtual Grantee Platform funded by BD², the BD² Hub (Notion), will be used as a Research Outputs Management System (ROMS). The ROMS will track resources and develop a catalog of research outputs to enhance access to these resources within the BD² network. All research outputs across all programs, including datasets, code, protocols, and resources, must be logged on the ROMS. Logging an output consists of sharing a link to the resource and annotating the entry with appropriate metadata.

- For resources that have been referenced in publications, a persistent identifier must be shared, such as a DOI or RRID.
- In the absence of a persistent identifier, such as in the case of sharing prior to publication, either a weblink to the resource or a file can be shared.

All members of the network will have access to the ROMS and can search the databases for resources.

Minimum Requirements:

- All research outputs generated through BD² funding must be catalogued on the BD² Hub no later than the time of publication of an associated BD² Manuscript.

Additional Recommendations

- Investigators should share outputs as early as possible and may add a resource to a catalog prior to making it publicly available.

Appendix A: Open Science Policy Checklist

The list below briefly outlines the core components of the BD² Open Science Policy that must be adhered to by all BD² Funded Investigators. Because most of these requirements are enforced at the time of publication, this section can be used as a checklist when preparing a manuscript. Detailed instructions on each item, as well as additional recommendations and guidelines, can be found in the full text of the policy.

- I. Preprint Requirement**
 - ☐ Published BD² Manuscripts with primary BD²-funded data must be preceded by a preprint.
 - ☐ Preprints must be deposited no later than the time of first submission to a journal.
 - ☐ Preprints must be open access and use a CC BY or CC BY 4.0 license (or a CC0 dedication).
 - ☐ Preprints must be submitted to BD² for an Open Science Check upon deposition.
- II. Manuscript Sharing Requirement**
 - ☐ Published BD² Manuscripts with primary BD²-funded data must be made immediately open access.
 - ☐ Published BD² Manuscripts must use a CC BY or CC BY 4.0 license (or a CC0 dedication).
 - ☐ Published BD² Manuscripts must include a Key Resources Table (KRT).
 - ☐ BD² Manuscripts must be submitted to BD² for an Open Science Check 30+ days before final publication.
- III. Acknowledgement Requirement**
 - ☐ Preprints, published manuscripts, and resource outputs must acknowledge BD² support.
 - ☐ Authors within the BD² Network must include BD² in their affiliations.
- IV. Method and Code Sharing Requirement**
 - ☐ All protocols used in the context of BD²-funded work must be published and open access.
 - ☐ All code used or developed in the context of BD²-funded work must be shared.
 - ☐ If code cannot be shared due to IP restrictions, BD² must be notified before as early as possible.
 - ☐ Persistent identifiers for all protocols and code must be cited in manuscript text and in the KRT.
- V. Data and Resource Sharing Requirement**
 - ☐ All BD²-funded datasets must meet FAIR Standards and deposited in an open access repository.
 - ☐ Human-derived datasets must have safeguards to protect sensitive information; PHI can never be shared.
 - ☐ All research tools generated with BD² funding must be catalogued in an appropriate repository.
 - ☐ All key laboratory resources used under BD² funding must be clearly identified with an RRID.
 - ☐ Persistent identifiers for all datasets and resources must be cited in manuscript text and in the KRT.
- VI. Network Sharing Requirement**
 - ☐ All resources generated (protocols, code, datasets, etc.) must be catalogued on the BD² Hub.
 - ☐ All resources generated must be made shareable within the network.

Appendix B: Endorsements and Glossary

Endorsements

This policy was developed using the [Open Science Policy from Aligning Science Across Parkinsons \(ASAP\)](#) as a framework. BD² additionally endorses the following organizations and efforts:

- ASAPbio Preprint Policy Framework – This framework outlines best practices in sharing preprints.
- cOAlition S - This coalition of organizations endorses Plan S, an initiative that promotes open access publication of research funded by public grants.
- FAIR Principles - These guidelines promote research data to be findable, accessible, interoperable, and reusable.
- Panton Principles - These principles promote open access to research data to allow for the analysis, processing, and sharing without financial, legal, and technical barriers.
- SfDORA - The Declaration on Research Assessment promotes eliminating journal-based metrics, assessing merit-based research rather than journal reputation, and improving publication requirements to improve research.

Glossary

Author Accepted Manuscript: The final version of a manuscript that has been accepted for journal publication, including all revisions, tables, graphics, and supplemental materials, but before final formatting or typesetting have been incorporated by the publisher.

BD² Data: Any dataset that was collected through an effort that was fully or partially funded by BD².

- **Integrated Network Data:** BD² Data collected from the Integrated Network study. This dataset consists of human subject data and has additional layers of protection.

BD² Biospecimens: Any biological materials collected from participants in a BD²-funded effort, including from the IN Longitudinal Cohort Study.

BD² Manuscript: Any manuscript that results from an effort that was fully or partially funded by BD², including any manuscript that includes primary BD² Data or the use of BD² Biospecimens.

- This definition includes manuscripts created after the funding period has ended.
- Manuscripts generated using shared institutional resources and/or core facilities that receive BD² funding are excluded unless the research project itself received direct BD² funding. Appropriate acknowledgment of core facilities' funding sources should follow institutional guidelines.
- Secondary manuscripts such as review articles or commentaries are only pursuant to the Manuscript Sharing Requirement (Section II) and the Acknowledgement Requirement (Section III).

Key Resources Table (KRT): A KRT lists the resources used and generated within a manuscript, including datasets, protocols, code, and laboratory materials such as cell lines, antibodies, viral vectors, and mouse lines. Each resource should be annotated across a standard set of fields that clearly identify the resource name, type, and source, as well as a persistent identifier and any usage details. A template is available for BD² investigators [here](#).

Licenses: BD² Manuscripts and resources must use a license that permits reuse. Licenses that restrict commercial use are not allowed, per the guidance from [cOAlition S](#). In selecting a license, authors should use the following guidelines:

- CC BY 4.0 (or CC BY) is the default license selection for research papers and outputs. This license ensures that authors receive attribution for the work but allows work to be freely shared (i.e., redistributed in any medium or format or adapted for future purposes).
- CC0 is a public domain dedication and has zero restrictions on re-use; under this license, no attribution is required. It is recommended for the sharing of datasets and a CC0-govt is required for the work of some US Federal Government employees.
- BD² does not allow CC-BY-ND or CC BY-NC-ND as licensing options for BD² Manuscripts. These licenses are not open-access licenses, as they do not permit adapting, transforming, translating, or updating the work.

Persistent Identifier: A persistent identifier is a unique and standardized code for an organization, individual, resource, or other digital object that is long-lasting. In contrast, a non-persistent identifier, such as a web address or a vendor catalog number, may not be maintained or may change over time, making a resource harder to locate or identify.

- Digital Object Identifier (DOI) – Used for articles, dataset uploads, and other basic files.
- Open Research and Contributor Identifier (ORCID) – Used for researchers.
- Research Organization Registry (ROR) – Used for organizations.
- Research Resource Identifier (RRID) – Used for lab resources such as cell lines, biosamples, and animal lines.

Appendix C: Best Practices in Resource Sharing

To ensure that all resources that are shared can benefit the research community, a set of best practices are recommended to guarantee that resources can be effectively leveraged and re-used. All resource sharing, whether it be protocols, datasets, code, etc., should incorporate the concepts FAIR Principles as follows:

- Findable – Clearly indexed with a persistent identifier.
- Accessible – Clear information on what the resource is and how it was generated, with metadata included.
- Interoperable – Non-proprietary and machine-readable format.
- Reusable – Least restrictive license applied and thoroughly described metadata are included.

Protocols: Each method used in a BD² Manuscript is required to be associated with a protocol on a public repository such as protocols.io or bioprotocols.

Protocols should:

- Use “recipe style” formatting, detailing step-by-step instructions and including “working” or “hidden” knowledge. **Anyone with relevant expertise should be able to repeat the protocol exactly.**
- Use clear identification of experimental resources through the citation of persistent identifiers, such as DOIs, Resource Research Identifiers (RRIDs), and Chemical Abstracts Service (CAS) numbers; supplier catalog numbers and URLs are not considered persistent identifiers.

Each protocol should represent exactly what was done for the manuscript in which it was cited. If a method is used in multiple publications, it need only be published once, though it should be versioned or forked as needed to reflect the experiments in an individual manuscript.

Code: All code used in a BD² manuscript must be shared on a public repository.

To promote reuse and transparency, each set of code should have a README file in the root directory where it is deposited. The README file should describe:

- The purpose of the code,
- The type of data the code is intended to analyze, process, or otherwise manipulate, including supported file types,
- The expected output of the code,
- Full documentation, including steps for running the code, system requirements, and installation, and
- License information and acknowledgements.

Curated examples of README files can be found [here](#) and several README file templates are available online, such as [here](#).

Datasets: Each dataset represented in a BD² Manuscript must be shared on a public repository.

Investigators are expected to share all raw and clean data using a format that conforms to field-specific recommendations. Because these recommendations can vary widely depending on the size and type of the data, each team must create a personalized data sharing plan at the outset of their project. Further, investigators are encouraged to meet with the BD² Program Team to ensure that the data sharing plan is in line with the requirements of the Open Science Policy. Resources are available on the BD² Hub to help investigators identify appropriate repositories.

General recommendations and expectations for all BD² Manuscripts are as follows:

- Human-subjects data must take necessary precautions to protect participant identity; PHI may never be shared.

- Many publications will have multiple types of datasets that will be deposited on separate repositories. At a minimum, this should include an upload to Zenodo that contains:
 - A README file with all metadata necessary to understand, assess, and replicate the reported study findings; a template is available for BD² investigators [here](#),
 - A Key Resources Table (KRT) for the BD² Manuscript, and
 - All basic tabular data, which should use [tidy data format](#). Tabular datasets can either be uploaded as individual spreadsheets or as one spreadsheet with an individual tab for each dataset.
 - Files should have clear labels such as '2025-Smith-Figure01a'.
 - Examples of basic tabular data include, data underlying a bar graph of participant age across two groups or a bar graph of protein concentration as measured by a Western blot.
- For all datasets across all types:
 - All formats should be standardized and machine-readable.
 - All formats should be accessible and non-proprietary.
 - All labels, including within the datasets and filenames, should be clear and descriptive.
 - All metadata necessary to understand how the data were collected should be included.
- Including a Data Availability Statement is recommended for all BD² Manuscripts. Example text as follows:
 - All data generated in this study are openly available and links to all datasets can be found in the Key Resources Table ([http://doi.org/\[doi\]](http://doi.org/[doi])).
 - All data generated in this study are openly available in [Repository Name] at [http://doi.org/\[doi\]](http://doi.org/[doi]).