

SHARED: Standards for High-Quality And Rigorous Explanation of Data

Dr. Alaina Pearce

Section	Description	Page Number	Table/Figure Number
Introduction	<input type="checkbox"/> Describe purpose of the dataset		
	<input type="checkbox"/> Describe relevant societal and/or scientific value of dataset <ul style="list-style-type: none"> <li>• Why is data needed?</li> <li>• What value does data provide beyond your primary research paper?</li> </ul>		
Related Resources	<input type="checkbox"/> Associated publications disclosed (e.g., research papers using the data)		
	<input type="checkbox"/> Companion project repository linked (if applicable)		
	<input type="checkbox"/> Companion code repository linked (if applicable)		
Methods: General	<input type="checkbox"/> Study design described		
	<input type="checkbox"/> Ethical approvals and protocols described		
Methods: Sample (for animal and human subjects)	<input type="checkbox"/> Description of participants/subjects		
	<input type="checkbox"/> (Human subjects) Participant characteristics table including descriptive statistics (if possible) overall and/or by group for: <ul style="list-style-type: none"> <li>• Age</li> <li>• Sex</li> <li>• Gender</li> <li>• Race</li> <li>• Ethnicity</li> <li>• Indicators of socioeconomic status and/or social determinants of health</li> </ul>		
	<input type="checkbox"/> Inclusion and exclusion criteria		
	<input type="checkbox"/> (Human subjects) Consent/assent procedures (human subjects only) including: <ul style="list-style-type: none"> <li>• Consent procedures for data reuse and sharing</li> </ul>		
	<input type="checkbox"/> (Human subjects) Describe any data access controls/limitations		
	<input type="checkbox"/> (Human subjects) Procedures/steps taken to ensure confidentiality is maintained		

Section	Description	Page Number	Table/Figure Number
Methods: Protocol	<input type="checkbox"/> Clear description of data collection protocols including: <ul style="list-style-type: none"> <li>• Dates</li> <li>• Locations</li> <li>• Equipment used</li> <li>• Software used</li> </ul>		
	<input type="checkbox"/> Variables measured and associated protocols including: <ul style="list-style-type: none"> <li>• Data acquisition protocols (e.g., task designs, sampling methods, etc.)</li> <li>• Rational for each variable</li> <li>• Relevant citations for methods/protocols</li> </ul>		
	<input type="checkbox"/> Describe metadata location/convention. Metadata should include: <ul style="list-style-type: none"> <li>• Variable names and definitions – if a naming convention was used, describe</li> <li>• Variable units and/or allowed values</li> <li>• Levels and definitions for categorical data</li> <li>• Whether the variable was derived (e.g., computed value based on other variables)</li> </ul>		
Methods: Data Processing	<input type="checkbox"/> Steps taken to get data in current (shared) state including software/code used to process data		
	<input type="checkbox"/> Missing data conventions <ul style="list-style-type: none"> <li>• Note: it is not recommended to include imputed data because each user should decide the best approach for dealing with missing data</li> </ul>		
	<input type="checkbox"/> Describe how low-quality data points were assessed for each variable (e.g., chance task performance, failed assay, etc.)		
Data Description: Data Access	<input type="checkbox"/> Repository name(s) and persistent identifier/DOI(s) <ul style="list-style-type: none"> <li>• Note: if multiple data repositories were needed, include how the data can be linked between repositories</li> </ul>		
	<input type="checkbox"/> License information		
	<input type="checkbox"/> Access conditions/controls for sensitive data		

Section	Description	Page Number	Table/Figure Number
Data Description: Overview	<input type="checkbox"/> General description of data deposited in repository including: <ul style="list-style-type: none"> <li>• Number of files</li> <li>• Size of files</li> <li>• File formats – open formats preferred</li> <li>• Metadata standard used</li> </ul>		
	<input type="checkbox"/> Structure of data in repository including: <ul style="list-style-type: none"> <li>• Directory structure</li> <li>• Links/relationships between data files</li> <li>• Naming conventions and/or data standard used</li> <li>• Important metadata files</li> </ul>		
Data Description: Key Variables	<input type="checkbox"/> Descriptive statistics about key variables such as: <ul style="list-style-type: none"> <li>• Observed value ranges</li> <li>• Central tendency and variability/variation</li> <li>• Number of missing values</li> <li>• Number of low quality/flagged values</li> </ul>		
Technical Validation/Quality Control	<input type="checkbox"/> Quality control metrics and/or protocols for each variable or set of variables (e.g., blinding, randomization, etc.)		
	<input type="checkbox"/> Validation of data (e.g., cross-checks, error analysis, signal quality)		
	<input type="checkbox"/> (if applicable) Phenotypic or genotypic assessments of biological samples (e.g., to confirm cell line or disease status)		
	<input type="checkbox"/> Known systemic issues or limitations to data		
	<input type="checkbox"/> Describe how sample-specific recommendations for inclusion/exclusion are documented (e.g., in a data file or in metadata)		

Section	Description	Page Number	Table/Figure Number
Discussion	<input type="checkbox"/> Describe the data's unique value (e.g., size, sample, resolution, diversity of measures), novel and/or expensive methods)		
	<input type="checkbox"/> Describe how the data fills a resource gap if one exists		
	<input type="checkbox"/> Describe how the data can be applied and/or reused: <ul style="list-style-type: none"> <li>• Possible research questions the data could help answer</li> <li>• If relevant, describe interdisciplinary uses for the data</li> <li>• If relevant, describe data standardization and note ability to integrate with other datasets</li> </ul>		
	<input type="checkbox"/> Describe limitations or potential pitfalls for users <ul style="list-style-type: none"> <li>• Any guidance or recommendations to address these when data is used</li> <li>• Warnings about out-of-scope analyses or potential misuse</li> </ul>		
Funding	<input type="checkbox"/> All funding sources with grant numbers, if applicable		