**BEAMRAD Tool for data users**

**Title**

1. **Please describe the title/acronym of the dataset.** If it is not available, please justify.
2. **Please describe how to cite the resource.** If no citation format is provided, please justify.

**General description**

1. **What are the target entities?** i.e., the anatomical region of subjects from whom the data was acquired (e.g., brain shown in MRI). Please describe them and if they can not be provided, please justify.
2. **What data type is included in this dataset?** Image data (e.g. MRI) , signal data (e.g., ECGs) or tabular data (e.g., EHR)? Please describe it or if it can not be provided, please justify.
3. **What is the principal motivation and objective for the dataset described?** If no information can be made available, please justify.
4. **Are there different versions of this dataset?** If so, please describe the principal motivation and objective for updating this dataset described? What is the current version?
5. **List the keywords for the dataset** (e.g., datatype, modality, body part, relevant medical terms).
6. **Provide the list of coordinators or organizers involved in the dataset creation?** If not available, please justify.
7. **Please specify the date of the first dataset publication.** If missing, please justify.
8. **Is the dataset associated with a challenge?** If so, what is the challenge name an official hosting place?

**Dataset usage**

1. **What is the official website or other place where the dataset is hosted?** What is the location (e.g., a URL).
2. **What is the accessibility of the dataset?** (choose from provided categories). Are there any specific barriers or formal agreements?
   1. Open access – no requirements for access or only sign-in
   2. Open access with barriers – datasets fulfilling the theoretical criteria for open access, but are inaccessible because of unpredictable reasons (e.g., no response to requests or broken hyperlinks)
   3. Regulated access – required the fulfilment of formal agreements or approvals.
3. **Please describe the ethics approval or reference to it. If not available, please justify.**
4. **Please describe the data usage agreement and licensing terms.** If no such information is present, please justify.
5. **Please describe the funding or sponsorship**. If not mentioned, please justify.
6. **Please describe all publication restrictions.** If not mentioned, please justify.
7. **Please describe the timeframe during which the data was acquired. (e.g., between 2005-2020)?** If this cannot be described, please justify.
8. **Please describe the origin of the data (e.g., country, institution).** If this cannot be specified, please justify.
9. **Please indicate whether this is a single-center or multi-center study.** If this cannot clarified, please justify.
10. **What are the criteria for data inclusion described? (e.g., consecutive participants, inclusion and exclusion criteria)**. If this cannot be clarified, please justify.
11. **What were the relevant characteristics (e.g., level of expertise) of the subjects (e.g., surgeon)/objects (e.g., robot) involved in the data acquisition process?** If not clarified, please justify.
12. **Please describe or refer to data anonymization protocol?** If this cannot be clarified, please justify.

**Metadata**

1. **Please describe participants’ metadata (e.g., age, medical history)? If relevant, what are the features, variable types, and the number of categories provided for the variable?**
2. **Please describe any accompanying data or metadata.**
3. **Please describe the hardware used for data acquisition.** (i.e. the vendor and the specific model)
4. **Please describe the data acquisition protocol.** If missing, please justify.

**Sample size**

1. **What is the number of participants and the number of data samples?** If this information cannot be made available, please justify
2. **If there is more than one inclusion for a participant, what is the number of data samples per participant?**

**Missing data**

1. **Please describe (and quantify) any information on missing data.** If none is available, please justify

**Training and test set distribution**

1. **Is data divided into training/test sets?** If the dataset consists of training/test sets, how is the data distributed?
2. **Are both the training and test datasets available?** If not, describe why.
3. **Please explain an potential differences between distributions of the data subjects in the training and test sets and the real-world distribution?** (e.g., data comes from different hospitals) If not addressed, justify.

**Data Annotation**

1. **What types of annotations are included in this dataset?**
2. **Please describe the annotation methods**. If not provided, please justify.
3. **Please describe if the human annotators were involved.** If no such info is available, please justify.
4. **If applicable, please describe or provide a reference to the annotation instruction protocol.** If not available, please justify.
5. **Please indicate the number of annotators involved.** If not specified, please justify.
6. **Please describe how annotators were selected.** If not addressed, please justify.
7. **Please describe annotators’ disciplinary background and level of expertise (e.g., in years).** If not mentioned, please justify.
8. **If multiple annotations are used, please describe the consensus approach.** If not available, please justify.

**Data preprocessing**

1. **Please describe the data preprocessing steps**. If none are provided, please justify

**Sources of errors**

1. **Please describe the most relevant possible sources of error (e.g., error through annotation variability).** If not documented, please justify.
2. **Please quantified the most relevant error sources.** (For instance, by giving specific information for training, validation and test cases?). If missing, please justify
3. **Please describe limitations of the dataset**