OND01 ALL 01 NIBS SYNTHDATA Methods

On behalf of the Neuroinformatics & Biostatistics platform

This methods document is organized into four sections:

1. Methods summary
2. Methods details
3. References
4. Contributions and contacts

Methods Summary

The synthetic data (SYNTHDATA) package is meant to serve as an illustration of the data standards and data package structure for ONDRI data, as well as illustrative data for data preparation and outlier detection. The synthetic data were created with the `synthpop` package in R (https://CRAN.R-project.org/package=synthpop), based on data from ONDRI. Missing data were introduced after the synthetic data were created. The SUBJECT, VISIT, SITE, and DATE information have been changed.

Methods Details

SYNTHDATA: The synthetic data were created with the `ctree` method in `synthpop`. The data were created based on data from ONDRI including the clinical, neuropsychology, genomics, and neuroimaging (structural; SABRE) platforms. The code used to create the data is available upon request (currently housed in Gitlab on BrainCODE).

SYNTHDATA: Tabular

The data are described as above, and distributed "tabular" data according to the standards used by ONDRI.

SYNTHDATA: Non-tabular

The data are described as above, and distributed "non-tabular" data according to the standards used by ONDRI. In this case, the files (i.e., those in DATAFILES) are actually empty text files; those files contain no data.

References

Nowok, B., Raab, G. M., & Dibben, C. (2016). synthpop: Bespoke creation of synthetic data in R. *Journal of statistical software*, *74*(11), 1-26.

Contributions and contacts

Creation and authorship this document is outlined below. There are four roles, three of which are defined by the CRediT system (https://casrai.org/credit/) with adaptations of the definitions as necessary, see also (CITE: https://www.pnas.org/content/115/11/2557):

- *Conceptualization*: Ideas; formulation or evolution of overarching goals and aims.
- *Writing - original draft*: Preparation, creation and/or presentation of the published work, specifically writing the initial draft.
- *Writing - review & editing*: Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision
- *Additional contributions*: Contributions to this document not otherwise captured in Conceptualization, Writing (draft), or Writing (reviews/edits).

Conceptualization, Writing (draft), and Writing (review/edit) are not exclusive from one another but are exclusive from Additional Contributions. If one appears in Contributions they do not appear in Conceptualization, Writing (draft), and Writing (review/edit) and vice versa.

Conceptualization	Writing (initial draft)	Writing (review/edit)	Additional Contributions
Derek Beaton	Derek Beaton	Kelly Sunderland	Jedid Ahn Stephen Arnott Malcolm Binns Vivian Chau Allison Dilliot Donna Kwan Roberto Lentini Logan Lim Paula McLaughlin Joel Ramirez Chris Scott Ningcan (Vonna) Sun Kelly Sunderland Brian Tan

For information on OND01 ALL 01 NIBS SYNTHDATA, please contact Derek Beaton <dbeaton@research.baycrest.org>.