Big NIBS data: a repository and pipeline for big data analyses in NIBS

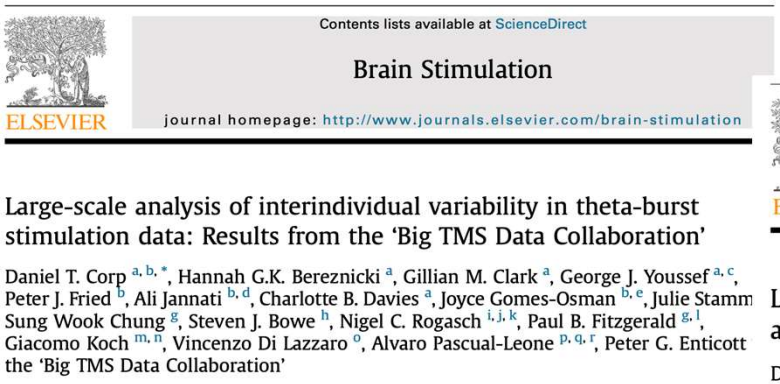
Daniel Corp, PhD

Deakin University, Australia

Outline

- Papers analysing motor evoked potential data
- Big NIBS data platform and repository
- All types of NIBS data. E.g. Clinical trials in PTSD
- How his evolves in the future
- NIBS-DAS (data analysis standard)

Big data analyses of MEP data (~2017-21)



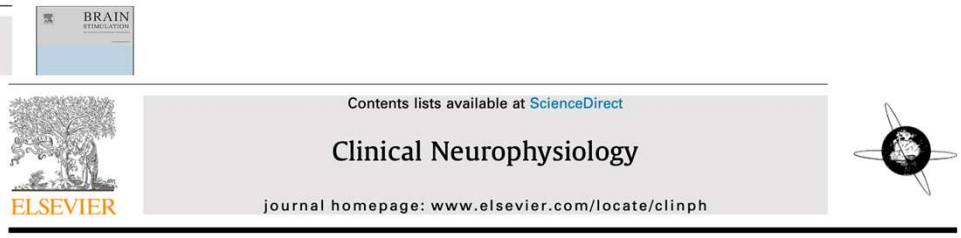
Contents lists available at ScienceDirect

Brain Stimulation

journal homepage: http://www.journals.elsevier.com/brain-stimulation

Large-scale analysis of interindividual variability in theta-burst stimulation data: Results from the 'Big TMS Data Collaboration'

Daniel T. Corp [a, b, *], Hannah G.K. Bereznicki [a], Gillian M. Clark [a], George J. Youssef [a, c], Peter J. Fried [b], Ali Jannati [b, d], Charlotte B. Davies [a], Joyce Gomes-Osman [b, e], Julie Stamm Sung Wook Chung [g], Steven J. Bowe [h], Nigel C. Rogasch [i, j, k], Paul B. Fitzgerald [g, l], Giacomo Koch [m, n], Vincenzo Di Lazzaro [o], Alvaro Pascual-Leone [p, q, r], Peter G. Enticott the 'Big TMS Data Collaboration'



Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Large-scale analysis of interindividual variability in single and paired-pulse TMS data

Daniel T. Corp [a,b,*], Hannah G.K. Bereznicki [a], Gillian M. Clark [a], George J. Youssef [a,c], Peter J. Fried [b], Ali Jannati [b,d], Charlotte B. Davies [a], Joyce Gomes-Osman [b,e], Melissa Kirkovski [a], Natalia Albein-Urios [a], Paul B. Fitzgerald [f,g], Giacomo Koch [h,i], Vincenzo Di Lazzaro [j], Alvaro Pascual-Leone [k,l,m], Peter G. Enticott [a], the 'Big TMS Data Collaboration'

- 118 corresponding authors of TMS studies were emailed and asked to provide deidentified individual TMS data.
- For theta-burst stimulation data, **22 studies, 430 healthy participants**
- For single/pp data, **35 studies, 687 healthy participants**
- Used mixed-effects regression to analyse factors driving response **inter-individual** variability

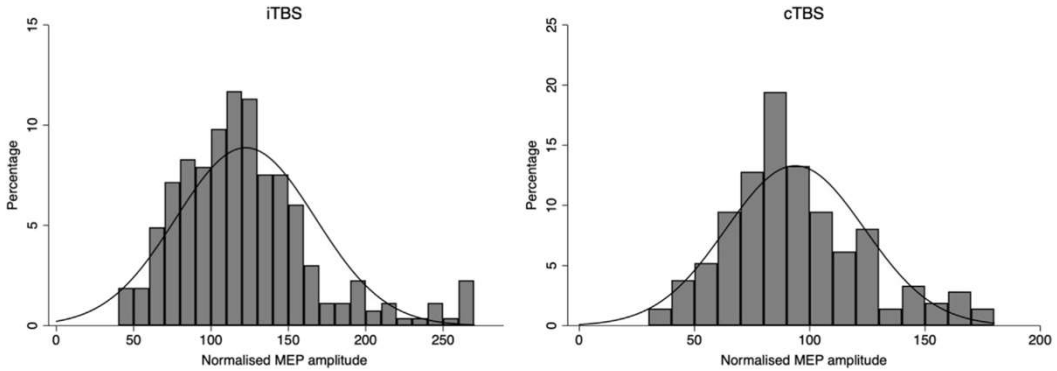
Big data analyses of MEP data (~2017-21)

- Must use correct statistical analyses (i.e., you can't just analyse data as though it all comes from one study).
- Mixed model regression. Random factors are participant and study ID
- These 'nest' data within participants and studies
- i.e., data from within participants/studies are more correlated than those from other participants/studies

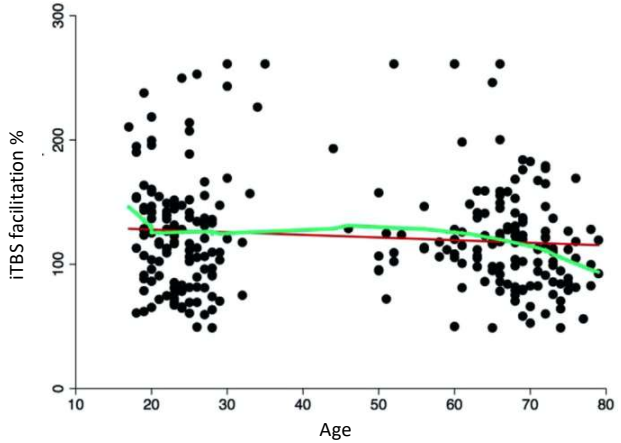
Corp et al., 2020, *Brain Stimulation*

Big data analyses of MEP data (~2017-21)

Some brief examples of TBS findings



1) No evidence of bimodal 'responders' and 'non-responders' (Lopez-Alonso et al., 2014; McAllister et al., 2013)



2) Non-linear relationship for iTBS response and age

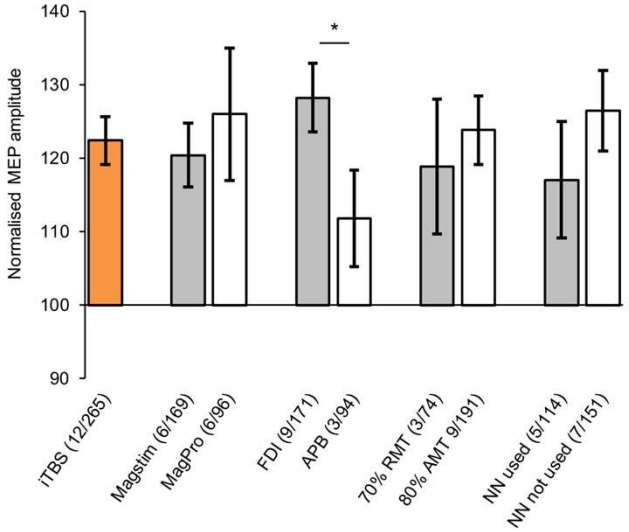
Hyper-excitability for age <20 and hypo-excitability for age >60.

Corp et al., 2020, *Brain Stimulation*

Big data analyses of MEP data (~2017-21)

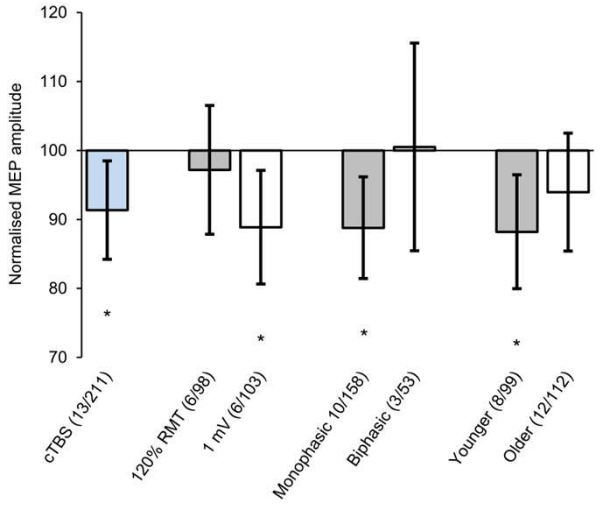
Some brief examples of TBS findings

iTBS



3) Significantly higher iTBS effect for FDI compared to APB

cTBS



4) cTBS effect when evaluated with monophasic, but not biphasic pulses

(Brackets in x-axes show N studies / participants for each IV level)

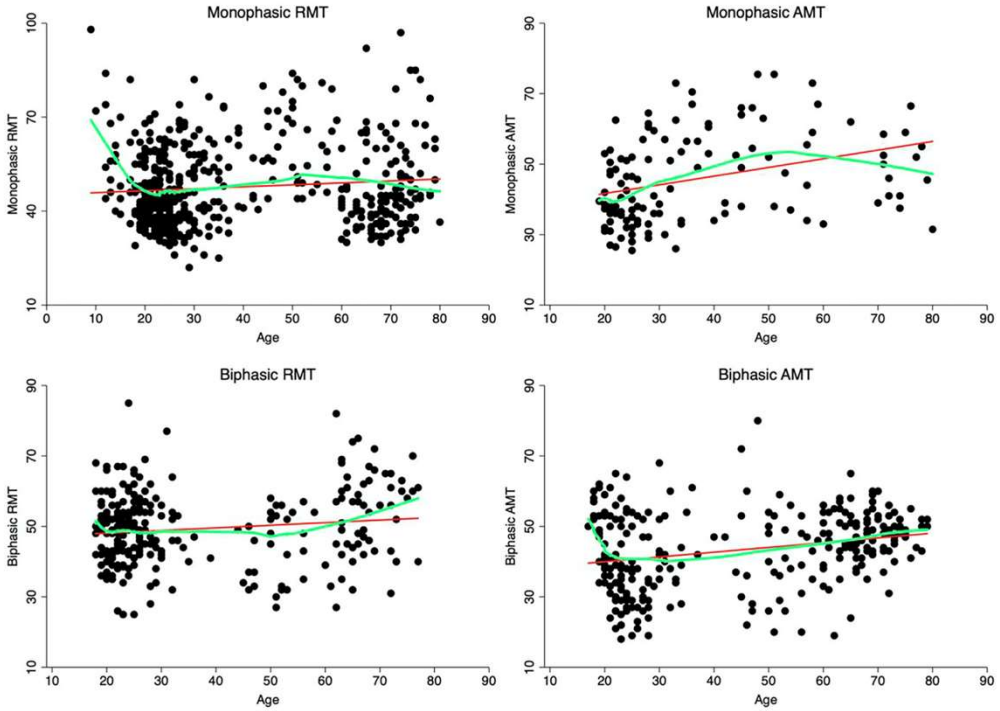
Corp et al., 2020, *Brain Stimulation*

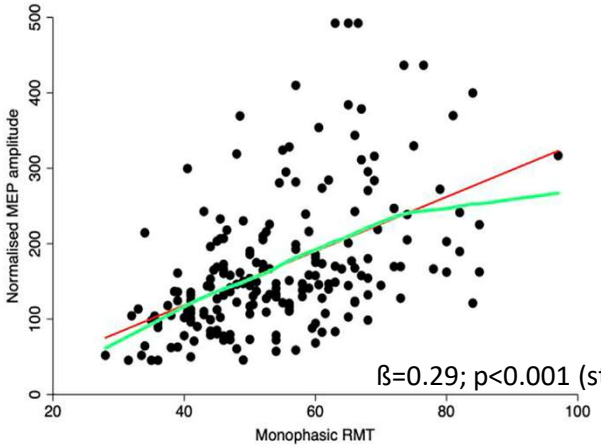
What have we done so far?

Examples of single/pp TMS findings

1) Non-linear relationships between MT and age.

Different patterns for mono vs biphasic pulses – likely due to different M1 micro-circuits activated by mono and biphasic pulses (Di Lazzaro et al,. 2001)





ß=0.29; p<0.001 (studies=11; N=193)

2) Significant relationship between MT and ICF (also sig relationship for SICI)

Corp et al., 2021, *Clinical Neurophysiology*

Big NIBS data platform and repository

- These results demonstrate that more data allowed us to identify factors that significantly predicted variability in TMS data → proves principle

However, its still quite 'low resolution' and summary level data:

- No real standardisation about how data were collected, managed, analysed, and formatted for sharing, between laboratories
- Insufficient metadata re methods and how data were collected meant we couldn't properly analyse these differences
- Mean data for MEP blocks; not individual MEPs
- % change scores

Big NIBS data platform and repository

Brain Stimulation 18 (2025) 278–279

Contents lists available at ScienceDirect

Brain Stimulation

journal homepage: www.journals.elsevier.com/brain-stimulation

Big non-invasive brain stimulation data (Big NIBS data): An open-access platform and repository for NIBS data

https://bignibsdata.com/

Upload data

Terms of use
Ensure that you agree to terms prior to sharing

Upload associated data
Upload non-NIBS data (e.g., MRI) to an existing repository (e.g., OpenNeuro). Then link to these data in our metadata

Format data
Format your data in our recommended structure (see website)

Provide metadata
Give details of experiment, type of data, links to associated data, and other details

Upload data
Submit your data to the repository

Validation and release
We will ensure that the data are safe and virus free, then release it to the repository

Download data

Terms of use
Ensure that you agree to terms prior to downloading

Search the repository
See our instructions for how to quickly search the repository to see whether it contains the data you want

Check list of pre-registered projects
Avoid duplication by ensuring your project is not already being done

Pre-register your project
If the project doesn't already exist, pre-register it and send the link to us

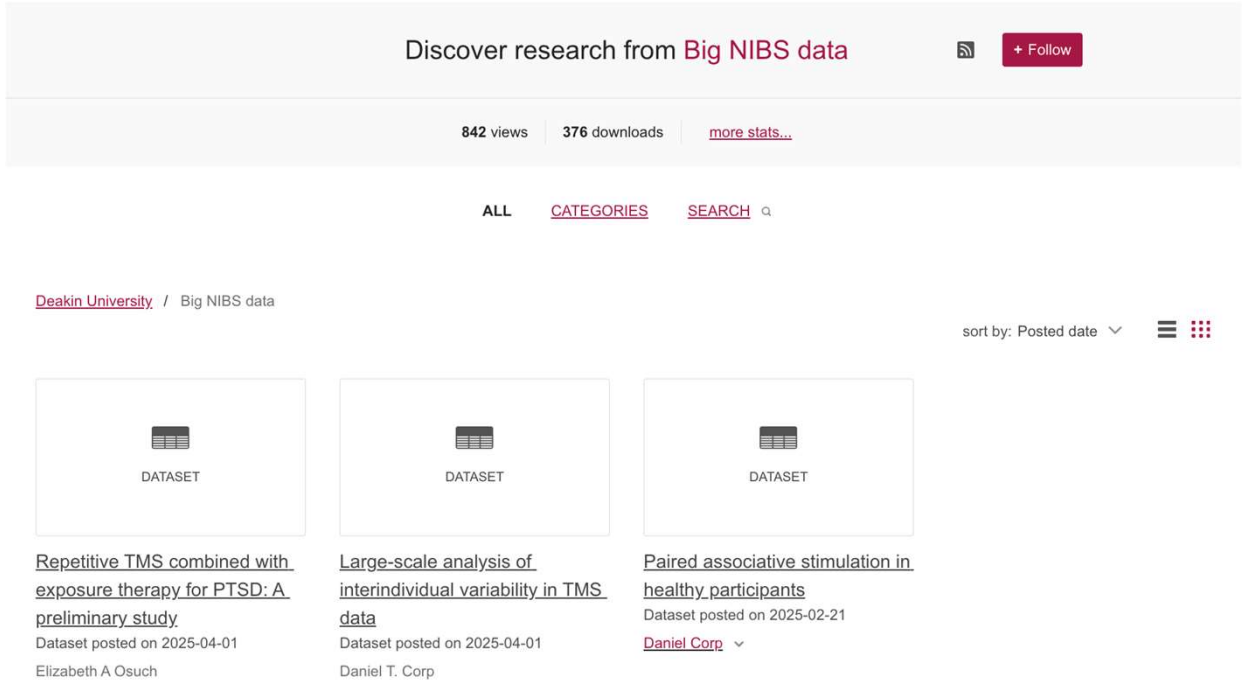
Download data
Download data from the repository

Download code
You are free to also download any code from the repository to help you handle your data for any NIBS project

Fig. 1. Overview of the data upload and download processes on the Big NIBS data repository.

Big NIBS data platform and repository

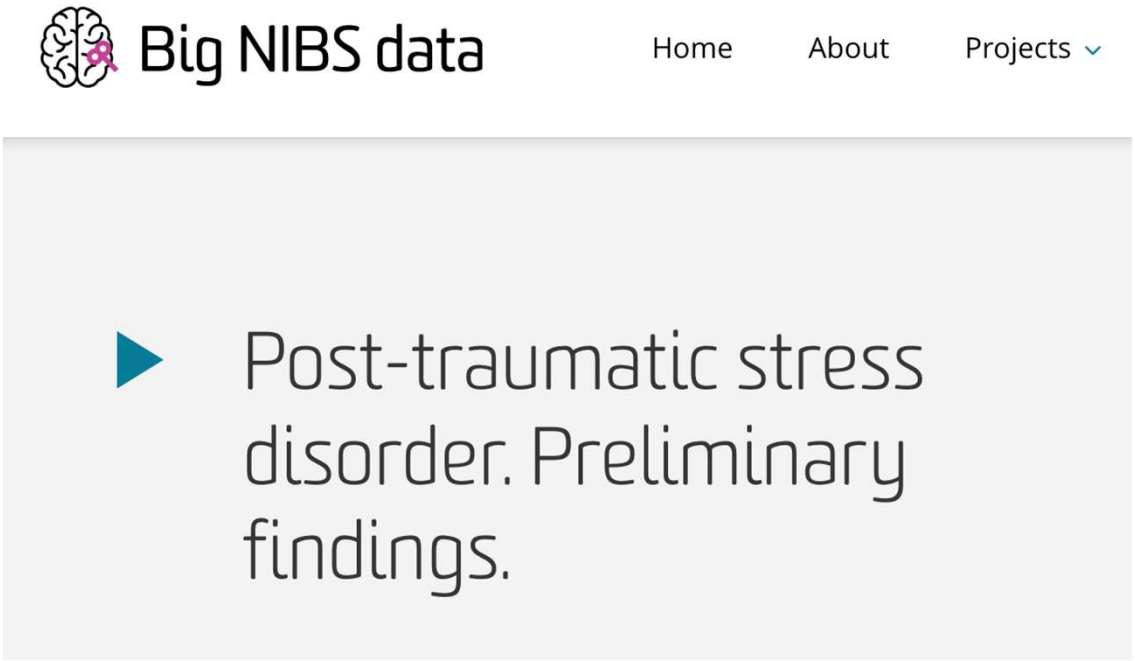
https://dro.deakin.edu.au/BigNIBSdata



Big NIBS data platform and repository

- The pipeline and repository are not perfect.
- We can move towards optimising these pipelines in the future

https://bignibsdata.com/completed-projects/



NIBS-DAS (data analysis standard)

- A Template for the Layout, Management, and Analysis of NIBS Data

- Unlike MRI, EEG, etc, this is the first repository and there are no standards for data sharing
- What is the format to share data to enable pooled analyses?
- Primary goal: providing a format for data upload to the Big NIBS data repository

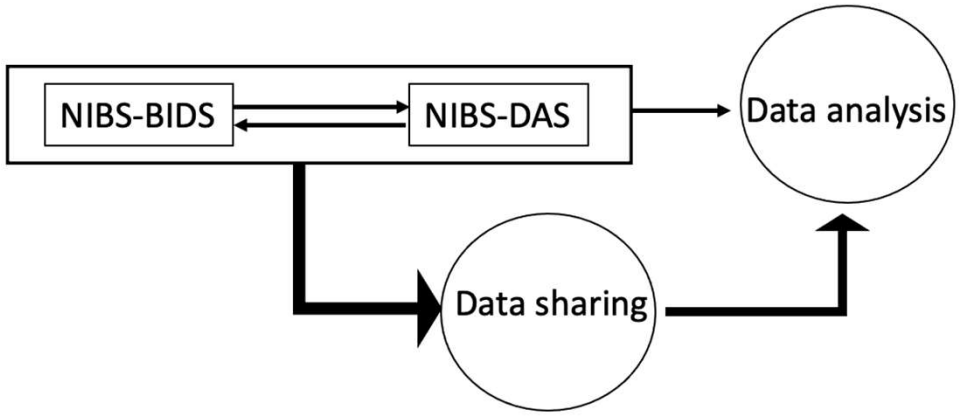


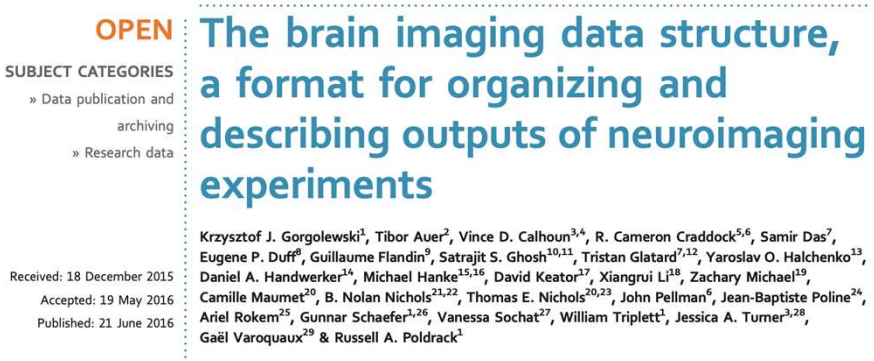
Figure 1. NIBS data management and analysis pipeline. The creation of NIBS-DAS for collated data allows for the integration with NIBS-BIDS (for raw data) to form a pipeline from data acquisition to data analysis, through to data sharing and re-analysis.

Michael Barham

NIBS-DAS (data analysis standard)

- This is complex – it must make sense, and there are existing structures for other modalities (e.g., MRI, EEG, etc).

- Secondary goal: help to standardise management and analysis practices
- These are extremely variable (and not very good IMO) in NIBS



OPEN

SUBJECT CATEGORIES

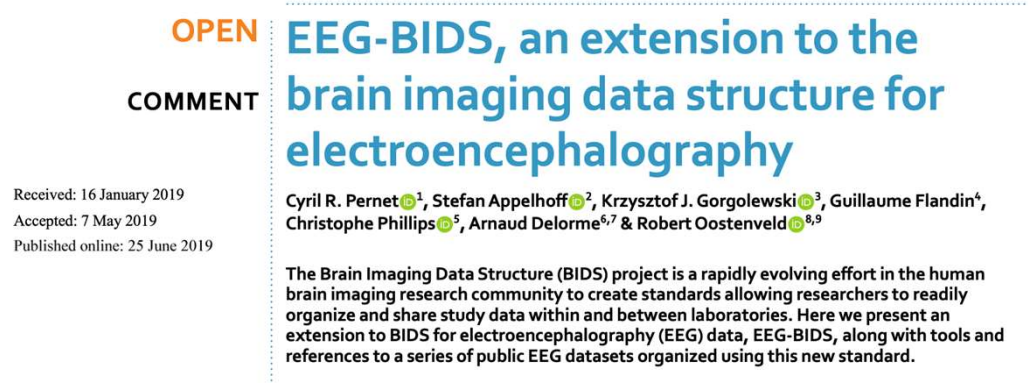
» Data publication and archiving

» Research data

The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments

Krzysztof J. Gorgolewski[1], Tibor Auer[2], Vince D. Calhoun[3,4], R. Cameron Craddock[5,6], Samir Das[7], Eugene P. Duff[8], Guillaume Flandin[9], Satrajit S. Ghosh[10,11], Tristan Glatard[7,12], Yaroslav O. Halchenko[13], Daniel A. Handwerker[14], Michael Hanke[15,16], David Keator[17], Xiangrui Li[18], Zachary Michael[19], Camille Maumet[20], B. Nolan Nichols[21,22], Thomas E. Nichols[20,23], John Pellman[6], Jean-Baptiste Poline[24], Ariel Rokem[25], Gunnar Schaefer[1,26], Vanessa Sochat[27], William Triplett[1], Jessica A. Turner[3,28], Gaël Varoquaux[29] & Russell A. Poldrack[1]

Received: 18 December 2015
Accepted: 19 May 2016
Published: 21 June 2016



OPEN

COMMENT

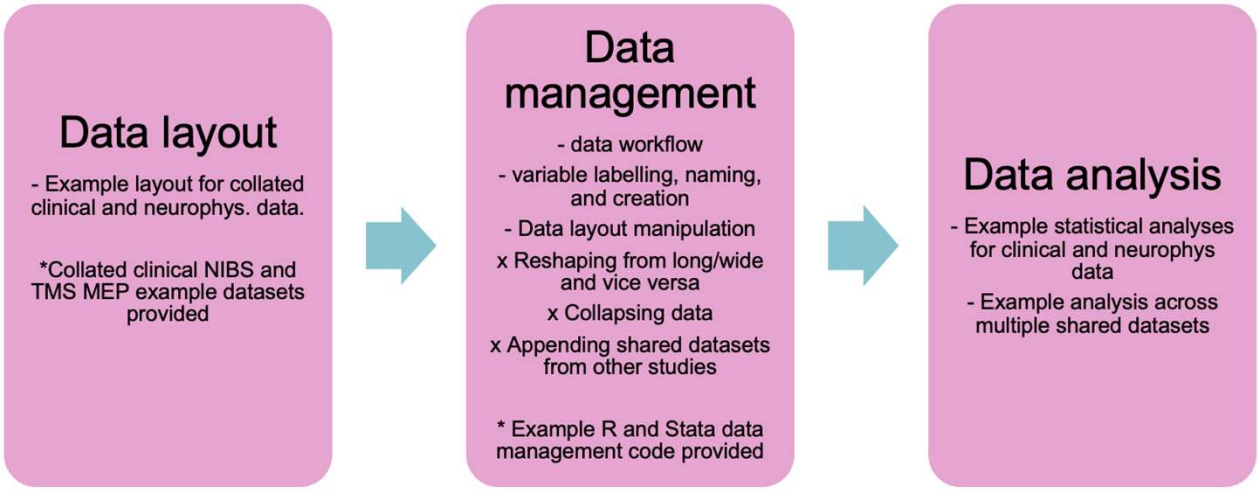
EEG-BIDS, an extension to the brain imaging data structure for electroencephalography

Cyril R. Pernet[1], Stefan Appelhoff[2], Krzysztof J. Gorgolewski[3], Guillaume Flandin[4], Christophe Phillips[5], Arnaud Delorme[6,7] & Robert Oostenveld[8,9]

Received: 16 January 2019
Accepted: 7 May 2019
Published online: 25 June 2019

The Brain Imaging Data Structure (BIDS) project is a rapidly evolving effort in the human brain imaging research community to create standards allowing researchers to readily organize and share study data within and between laboratories. Here we present an extension to BIDS for electroencephalography (EEG) data, EEG-BIDS, along with tools and references to a series of public EEG datasets organized using this new standard.

NIBS-DAS (data analysis standard)



Overall pipeline

Figure 3. The three components of NIBS-DAS: data layout, data management, and data analysis.

How data and metadata can be uploaded. Ideally…

This would be your project folder with ‘raw’ and collated data

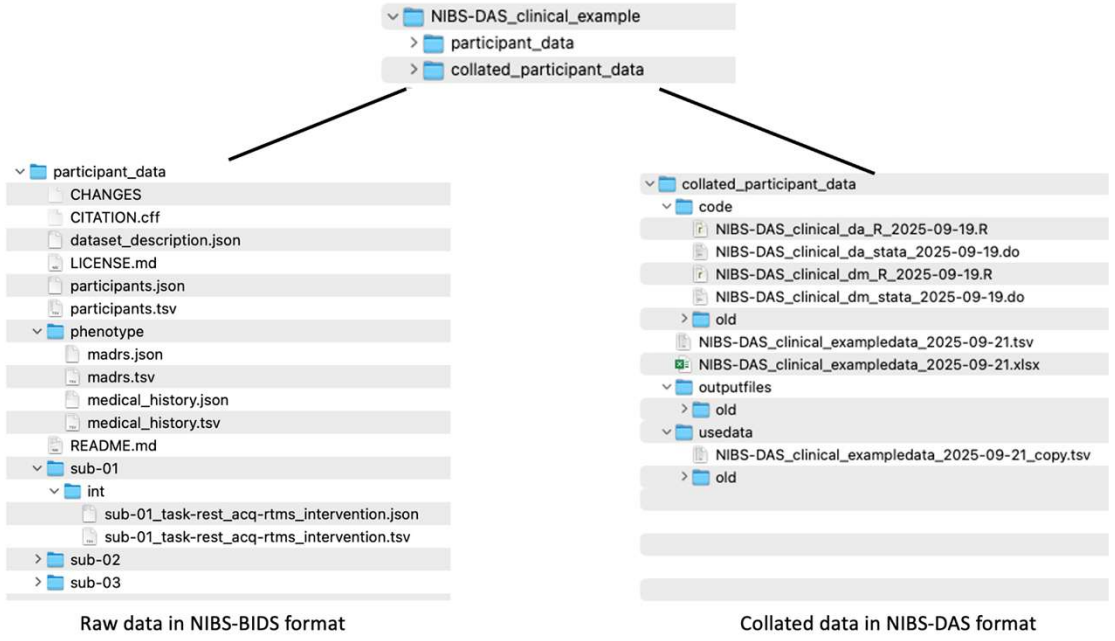
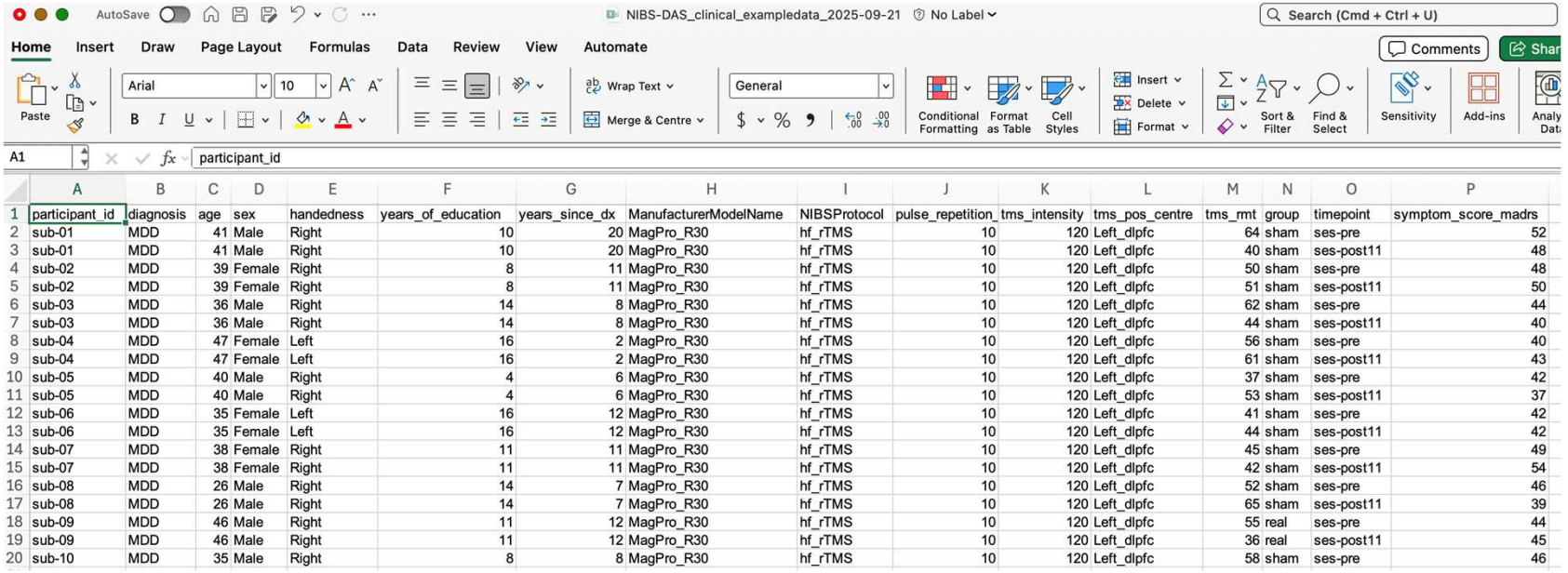


Figure 2. Example data structure for NIBS-DAS (right, focus of present article), and also an example of current guidelines for how the raw data can be structured in NIBS-BIDS format (left). These two structures can exist under the overarching project folder to store and utilise all raw and collated data from the project.

NIBS-DAS (data analysis standard)

1. Data layout



NIBS-DAS_clinical_exampledata_2025-09-21

	participant_id	diagnosis	age	sex	handedness	years_of_education	years_since_dx	ManufacturerModelName	NIBSProtocol	pulse_repetition	tms_intensity	tms_pos_centre	tms_rmt	group	timepoint	symptom_score_madrs
1	participant_id	diagnosis	age	sex	handedness	years_of_education	years_since_dx	ManufacturerModelName	NIBSProtocol	pulse_repetition	tms_intensity	tms_pos_centre	tms_rmt	group	timepoint	symptom_score_madrs
2	sub-01	MDD	41	Male	Right	10	20	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	64	sham	ses-pre	52
3	sub-01	MDD	41	Male	Right	10	20	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	40	sham	ses-post11	48
4	sub-02	MDD	39	Female	Right	8	11	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	50	sham	ses-pre	48
5	sub-02	MDD	39	Female	Right	8	11	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	51	sham	ses-post11	50
6	sub-03	MDD	36	Male	Right	14	8	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	62	sham	ses-pre	44
7	sub-03	MDD	36	Male	Right	14	8	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	44	sham	ses-post11	40
8	sub-04	MDD	47	Female	Left	16	2	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	56	sham	ses-pre	40
9	sub-04	MDD	47	Female	Left	16	2	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	61	sham	ses-post11	43
10	sub-05	MDD	40	Male	Right	4	6	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	37	sham	ses-pre	42
11	sub-05	MDD	40	Male	Right	4	6	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	53	sham	ses-post11	37
12	sub-06	MDD	35	Female	Left	16	12	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	41	sham	ses-pre	42
13	sub-06	MDD	35	Female	Left	16	12	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	44	sham	ses-post11	42
14	sub-07	MDD	38	Female	Right	11	11	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	45	sham	ses-pre	49
15	sub-07	MDD	38	Female	Right	11	11	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	42	sham	ses-post11	54
16	sub-08	MDD	26	Male	Right	14	7	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	52	sham	ses-pre	46
17	sub-08	MDD	26	Male	Right	14	7	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	65	sham	ses-post11	39
18	sub-09	MDD	46	Male	Right	11	12	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	55	real	ses-pre	44
19	sub-09	MDD	46	Male	Right	11	12	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	36	real	ses-post11	45
20	sub-10	MDD	35	Male	Right	8	8	MagPro_R30	hf_rTMS	10	120	Left_dlpfc	58	sham	ses-pre	46

2. Data management

NIBS-DAS_clinical_dm_stata_2025-09-25.do

```
/*******************************************************************************
Do-file title   NIBS-DAS_Clinical_DM_Stata.do

Stata code file demonstrating principles of data management (DM) of a fictional
clinical NIBS dataset: "NIBS-DAS_Clinical_ExampleData_Copy_YYYYMMDD.xlsx".

Code demonstrates (i) importing data to Stata, (ii) variable labelling, naming,
and creation, (iii) data checking, (iv) data layout manipulation (e.g., reshape
long and wide), and (v) appending multiple datasets.

*******************************************************************************/

clear all
capture log close
set more off
*version 17.0
version 15.0

*******************************************************************************
* OPEN DATA *
*******************************************************************************

* Set 'UseData' directory
*DC
cd "/Users/danielcorp/Dropbox/2016_BigTMSdata/0_Projects/0_2023_NIBS-DAS/NIBS-BIDS/FolderStructure_19.8.25/NIBS-DAS
* Open dataset
import excel using "NIBS-DAS_clinical_exampledata_2025-09-21_copy.xlsx", first

*DC. THis would be to import tsv but guess we can just to Excel.
*import delimited using "NIBS-DAS_clinical_exampledata_2025-09-21_copy.tsv", delimiters("\t") varnames(1) clear

*cd "/[YourFilePath]/UseData"
* Open dataset
*import excel using "NIBS-DAS_Clinical_ExampleData_Copy_YYYYMMDD.xlsx", first
*DC: Mike it should always be date_copy not other way around.

*******************************************************************************
* ASSIGN VALUES, ADD LABELS TO VARIABLES, AND CHECK DATA *
*******************************************************************************
```

3. Data analysis

NIBS-DAS_clinical_da_stata_2025-09-25.do

```
*******************************************************************************
* OPEN DATA *
*******************************************************************************

* Set 'UseData' directory
cd "/[YourFilePath]/UseData"

*******************************************************************************
* EXAMPLE I - ANALYSIS OF SYMPTOM SCORES BETWEEN CONDITIONS *
*******************************************************************************
/* This analysis requires a 'long' dataset layout. */

* Open file
use "NIBS-DAS_Clinical_ExampleData_Long_YYYYMMDD.dta", clear

* Clinical Example Question 1: is there a significant difference in pre and post TMS symptom scores between active and sham
mixed c.SymptomScore_MADRS c.Age i.Sex i.Timepoint##i.Condition || Participant_id:, base noretable
contrast i.Timepoint##i.Condition // Significance of overall interaction effect.
margins Condition, dydx(i.Timepoint)
margins i.Timepoint#i.Condition  // Show marginal means for each level of interaction between timepoint and condtion
marginsplot, ytitle("MADRS symptom scores")
cd "/[YourFilePath]/OutputFiles" // Change directory to 'OutputFiles' folder
graph export "Clinical_Question1_Figure_YYYYMMDD.tif", replace // Generate and export graph
```

Thank you!

Questions?

daniel.corp@deakin.edu.au