Pernet, Cyril; Garrido, Marta I.; Gramfort, Alexandre; Maurits, Natasha; Michel, Christoph M.; Pang, Elizabeth; Salmelin, Riitta; Schoffelen, Jan Mathijs; Valdes-Sosa, Pedro A.; Puce, Aina

Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research

Published in:
Nature Neuroscience

DOI:
10.1038/s41593-020-00709-0

Published: 01/12/2020

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Pernet, C., Garrido, M. I., Gramfort, A., Maurits, N., Michel, C. M., Pang, E., Salmelin, R., Schoffelen, J. M., Valdes-Sosa, P. A., & Puce, A. (2020). Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research. *Nature Neuroscience*, *23*(12), 1473-1483. https://doi.org/10.1038/s41593-020-00709-0

Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research

Authors – copy
Published in Nature Neuroscience
https://doi.org/10.1038/s41593-020-00709-0

Cyril Pernet, Marta I. Garrido, Alexandre Gramfort, Natasha Maurits, Christoph M. Michel, Elizabeth Pang, Riitta Salmelin, Jan Mathijs Schoffelen, Pedro A. Valdes-Sosa & Aina Puce

Dr Cyril Pernet (Co-Chair), Centre for Clinical Brain Sciences, The University of Edinburgh, Chancellor's Building, 49 Little France Crescent, Edinburgh EH16 4SB
cyril.pernet@ed.ac.uk ORCID 0000-0003-4010-4632

Marta I. Garrido, Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Australia. marta.garrido@unimelb.edu.au ORCID 0000-0003-0679-4959

Alexandre Gramfort, Université Paris-Saclay, Inria, CEA, Palaiseau France. alexandre.gramfort@inria.fr ORCID 0000-0001-9791-4404

Natasha Maurits, University of Groningen, University Medical Center Groningen, Department of Neurology, The Netherlands. n.m.maurits@umcg.nl. ORCID 0000-0001-6231-7044

Christoph M. Michel, Department of Basic Neurosciences, University of Geneva, & Center for Biomedical Imaging (CIBM), Campus Biotech, 9, Chemin des Mines, 1202 Geneva, Switzerland.
christoph.michel@unige.ch ORCID 0000-0001-6523-9635

Elizabeth W. Pang, PhD, Neurology / Neuroscience & Mental Health, Hospital for Sick Children / SickKids Research Institute, 555 University Avenue, Toronto, Ontario, Canada, M4G 1E9.
elizabeth.pang@sickkids.ca ORCID 0000-0001-5541-3329

Riitta Salmelin, Department of Neuroscience and Biomedical Engineering, Aalto University, PO Box 12200, FI-00076 AALTO, Finland. riitta.salmelin@aalto.fi ORCID 0000-0003-2499-193X

Jan Mathijs Schoffelen, Radboud University, Donders Institute for Brain, Cognition and Behaviour, P.O.Box 9104, 6500 HB, Nijmegen, The Netherlands. jan.schoffelen@donders.ru.nl ORCID 0000-0003-0923-6610

Pedro A. Valdes-Sosa, Joint China-Cuba Laboratory for Neurotechnology. The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Lab for Neuroinformation, University of Electronic Science and Technology of China, Chengdu, China & Cuban Neuroscience Center, Havana, Cuba.
pedro.valdes@neuroinformatics-collaboratory.org ORCID 0000-0001-5485-2661

Aina Puce (Co-Chair), Department of Psychological & Brain Sciences, Indiana University, 1101 East Tenth St, Bloomington IN 47401, USA.
ainapuce@indiana.edu ORCID 0000-0002-0716-4185

Abstract

The Organization for Human Brain Mapping (OHBM) has been active in advocating for the instantiation of best practices in neuroimaging data acquisition, analysis, reporting, and sharing of both data and analysis code, to deal with issues in science related to reproducibility and replicability. Here we summarize recommendations for such practices in magnetoencephalographic (MEG) and electroencephalographic (EEG) research, recently developed by the OHBM neuroimaging community known by the abbreviated name of COBIDAS MEEG. We discuss rationale for the guidelines and their general content, which encompasses many topics under active discussion in the field. We highlight future opportunities and challenges to maximizing the sharing and exploitation of MEG and EEG data, and also how this 'living' set of guidelines will evolve to continually address new developments in neurophysiological assessment methods and multimodal integration of neurophysiological data with other data types.

Keywords

best practices; data acquisition; data analysis; data sharing; magnetoencephalography; electroencephalography; COBIDAS; Organization for Human Brain Mapping

3997 words, 3 figures, 2 boxes, 3 tables, 88 references

The OHBM COBIDAS MEEG report

The neuroimaging community, like many other scientific communities, is actively engaged in open science practices designed to improve reproducibility and replicability[1] of scientific findings. The Organization for Human Brain Mapping (OHBM), through Committees on Best Practices in Data Analysis and Sharing (COBIDAS), promotes and distributes commonly agreed practices formalizing their terminology, in consensus with other organizations. OHBM has developed the COBIDAS reports[2,3] to present best practices for specific neuroimaging methods, propose a standardized scientific language for reporting and promote effective sharing of data and methods. The reports are useful to (i) those preparing manuscripts and grant proposals of their work, (ii) editors and reviewers, (iii) neuroimaging educators; and (iv) those with expertise in a neuroimaging technique who seek to become *au fait* with another.

In this Perspective, we focus on the COBIDAS MEEG report[2] highlighting some of the main issues and ensuing recommendations generated by the committee. Our purpose is to provide a better understanding of how some acquisition parameters, design, analysis and reporting choices *can influence reproducibility*. Beyond these, many other issues have also found their way in the recommendations (see boxes 1, 2 & tables 1, 2, 3). As such, these recommendations represent the *minimal requirements* to be reported to ensure reproducible MEEG studies, and for each recommendation full details can be found in the COBIDAS report itself[2]. At the same time, many of these seemingly basic pieces of advice are contentious. A great deal of discussion has been spent on terminology, and our proposal is a consensus that adopts and extends the terminology used in the Brain Imaging Data Structure (BIDS https://bids.neuroimaging.io/) that enables better data sharing (initially for MRI[4] and now also for neurophysiological data with MEG-BIDS[5], EEG-BIDS[6] and iEEG-BIDS[7]). It also follows nomenclatures of the International Federation for Clinical Neurophysiology (IFCN https://www.ifcn.info/) current clinical guidelines, thus integrating research and clinical practices. It is also clear to us that there is no best analysis workflow (even if some general principles exist, e.g. Fig. 2) or best statistical approach, only optimal solutions to a given problem - and *this is why reporting context, acquisition and analysis details are so important*.

The MEEG community has always been proactive in discussing good practices and reporting, evidenced by the long history of published guidelines[8–15]. Some aspects of these guidelines have remained current despite the rapidly changing developments in MEEG hardware/software and methods. While the OHBM COBIDAS MEEG report follows this tradition, it differs from previous guidelines in three important respects. First, it has a focus on practices that specifically aid with reproducibility and data sharing. Second, the COBIDAS MEEG report exists as a *living document* in the format of a WordPress blog that invites feedback and comments (https://cobidasmeeg.wordpress.com/), with version controlled preprint releases on the Open Science Framework (https://osf.io/a8dhx/). We invite readers to refer to this document[2] when preparing scientific material. There has been exponential growth in the MEG and EEG literature in the 21st century (see Fig. 1a). A dynamic guideline is important as there have been many updates of acquisition and analysis methods, and the implementation of new technologies needs also to be integrated while keeping a coherent

set of recommendations. For instance, portable EEG devices, portable MEG devices operating at room temperature, and Brain Computer Interfaces (BCI) have not been considered as these are still emerging technologies (Fig. 1 b,c). Yet as these become more extensively used and available, experience will grow and best practices will need development. Additionally, COBIDAS MEEG has not considered invasive EEG (iEEG) recordings, despite their long history and recent renewed interest. In future, these might be integrated under a more general 'COBIDAS Neurophysiology' document. Third, the target population for the COBIDAS MEEG guidelines is considerably broader and larger than that served by previous guidelines, which traditionally were targeted to members of neurophysiological societies or interest groups concerned with one specific imaging modality (EEG or MEG), analytical method (ERP, spectrum, source, etc) or practice (research or clinic).

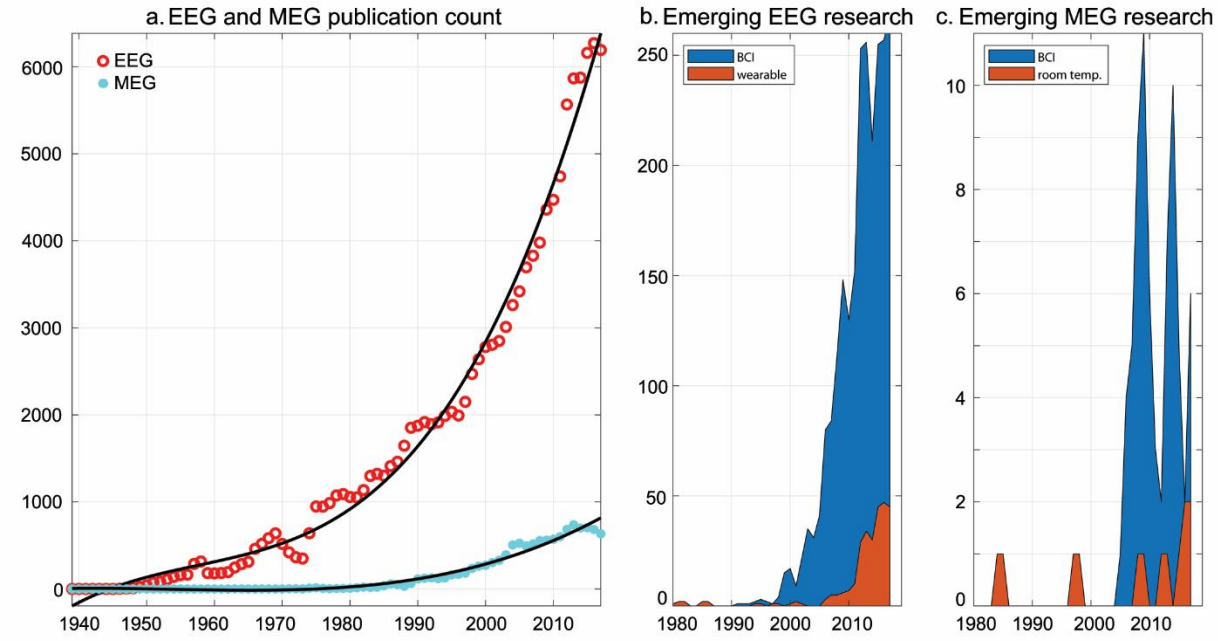


Figure 1. Overview of the total number of MEEG publications with emerging research fields a. *Number of EEG and MEG publications by year of publication.* ***b.*** *Emerging EEG research. Number of publications under the topics of Brain Computer Interface [BCI] and mobile/wearable EEG by year.* ***c.*** *Emerging MEG research. Number of publications by year for BCI and room temperature [Optically Pumped Magnetometer based] portable MEG. Source for literature searches: Medline.*

Terminology and reporting recommendations

To promote reproducible experimentation, one must share a common language. Some terms are common across imaging modalities, but can have slightly different usages. The COBIDAS MEEG terminology for describing task parameters and data acquisition follows those of the COBIDAS MRI and Brain Imaging Data Structure (Box 1). Of particular interest to MEEG researchers, we recommend using 'run' rather than 'block', which are used interchangeably in MEEG, but clearly differ for PET or MRI. Also, we recommend

explicitly reporting the space in which data processing (i.e. statistical analyses and modeling) is taking place: sensor vs. source. This is important as certain analytical methods may not be suitable for use in sensor space. While other data spaces have been reported in the literature, e.g. independent component space, these are only mathematical subspaces of the more general categories mentioned here.

There is also a specific MEEG terminology to describe features in the data that does not exist for MRI-based studies. Our recommendations (Box 2) are to follow conventions and common nomenclature[16], consistent with IFCN guidelines. We propose additional considerations for reporting EEG results aimed at reducing confusion in the literature as follows: (1) for reporting evoked data in sensor space, recording site(s) should be noted (e.g., vertex N100), as response polarity can vary by either original or post-hoc scalp reference electrode and underlying cortical folding; (2) latency windows used to quantify event-related components should be explicitly mentioned. For reporting spontaneous or resting-state MEEG data, in particular for spectral analyses, we advocate explicitly reporting boundaries of different frequency bands. There is confusion in the literature caused by inconsistencies in designating 'canonical' frequency bands[14,17] (e.g., delta, theta, alpha, beta, gamma). Here, we considered IFCN guidelines[14] for delineating canonical MEEG frequency bands, as these remain close to those originally proposed in the late 1920s by Berger[18], and in the 1930s by Walter[19], as well as Jasper and Andrews[16], and align with the main clinical textbook in the field[20]. That said, due to inconsistencies across literatures, we made a slight adjustment to the transition between alpha and beta ranges to guide results description for time-frequency analyses.

Session. A logical grouping of neuroimaging and behavioural data collected consistently across participants. A session includes the time involved in completing all experimental tasks. This begins when a participant enters the research environment and continues until he/she leaves. This would typically start with informed consent procedures, followed by participant preparation (i.e., electrode placement and impedance check for EEG; fiducial and other sensor placement for MEG). It would end when the electrodes are removed (for EEG) or the participant exits the MEG room, but could potentially also include a number of pre- or post-MEEG observations and measurements (e.g., anatomical MRI, additional behavioural or clinical testing, questionnaires), even on different days. Defining multiple sessions is appropriate when several identical or similar data acquisitions are planned and performed on all (or most) participants, often in the case of some intervention between sessions (e.g., training or therapeutics) or for longitudinal studies.

Run. An uninterrupted period of continuous data acquisition without operator involvement. Note that continuous data need not be saved continuously; in some paradigms, especially with long inter-trial intervals, only a segment of the data (before and after the stimulus of interest) are saved. In the MEEG literature, this is also sometimes referred to as a block. (Note the difference with the 'block' term in COBIDAS MRI, where multiple stimuli in one condition can be presented over a prolonged and continuous period of time.)

Event. An isolated occurrence of a presented stimulus, or a subject response recorded during a task. In addition to the identity of the events, it is essential to have exact timing information synchronized to the MEEG signals. For this, a digital trigger channel with specific marker values, or a text file with marker

values and timing information can be used. (This term has been defined here in a more narrow and explicit sense than that for COBIDAS MRI, mainly because of the specialized requirements surrounding the high temporal resolution acquisition of MEEG data.)

Trial. A period of time that includes a sequence of one or more events with a prescribed order and timing, which is the basic, repeating element of an experiment. For example, a trial may consist of a cue followed after some time by a stimulus, followed by a response, followed by feedback. An *experimental condition* is a functional unit defined by the design and usually includes many trials of the same type. Critical events within trials are usually represented as time-stamps or "triggers" stored in the MEEG data file, or documented in a marker file.

Epoch. In the MEEG literature, the term *epoch* designates the outcome of a data segmentation process. Typically, epochs in event-related designs (for analysis of event-related potentials or event-related spectral perturbations) are time-locked to a particular event (such as a stimulus or a response). Epochs can also include an entire trial, made up of multiple events to suit the data analysis plan. (This terminology is not used in the COBIDAS MRI specification.)

Sensors. Sensors are the physical objects or transducers that are used to perform the analogue recording, i.e., EEG electrodes and MEG magnetometers/gradiometers. Sensors are connected to amplifiers, which not only amplify, but also filter the MEEG activity.

Channels. Channels refer to the digital signals that have been recorded by the amplifiers. It is thus important to distinguish them from sensors. A 'bad channel' refers to a channel that is producing a consistently artifactual or low-quality signal.

Fiducials. Fiducials are markers placed within a well-defined location, which are used to facilitate the localization and co-registration of sensors with other spatial data (e.g., the participant's own anatomical MRI image, an anatomical MRI template or a spherical model). Some examples are vitamin-E markers, reflective disks, felt-tip marker dots placed on the face, or sometimes even the EEG electrodes themselves. Fiducials are typically placed at a known location relative to, or overlying, anatomical landmarks.

Anatomical landmarks. These are well-known, easily identifiable physical locations on the head (e.g., nasion at the bridge of the nose; inion at the bony protrusion on the midline occipital scalp) acknowledged to be of practical use in the field. Fiducials are typically placed at anatomical landmarks to aid localization of sensors relative to geometric data.

Sensor space. Sensor space refers to a representation of the MEEG data at the level of the original sensors, where each of the signals maps onto the spatial location of one of the sensors.

Source space. Source space refers to MEEG data reconstructed at the level of inferred neural sources that presumably gave rise to the measured signals (according to an assumed biophysical model). Each signal maps onto a spatial location that is readily interpretable in relation to the individual, or a template-based, brain anatomy.

99 ***Box 1. Specific MEEG terminology and definitions with respect to data acquisition.***
100

Event-related response component vs deflection. For time domain MEEG data, "component" traditionally refers to a functional brain process that has a characteristic spatial distribution and canonical latency[8]. Because of this loaded meaning for the term "component", the term "deflection" is a useful alternative.

Event-related response nomenclature. For EEG, event-related response components are named using a convention, where (EEG) response polarity and its *nominal* latency form the name (e.g., N100, N170, P300, N400, etc.), preferably adding the recording site. This was first published in the International Federation for Clinical Neurophysiology (IFCN) guidelines in 1983 (and updated in 1999), and advocated for reporting of clinical data[11], based on original nomenclature[8]. For MEG, the analogous components are referred to by two conventions: (1) an "m" added to the component name (e.g., N100m, N170m) or (2) referred to as M100, M170, etc.

Specialized MEEG event-related component nomenclature. Certain MEEG responses e.g. mismatch negativity (MMN), contingent negative variation (CNV), error-related negativity (ERN), among others, refer to specific responses elicited in particular types of paradigm, or to presumed mental states (e.g., error detection).

Other nomenclature. Early studies often refer to event-related components by successive EEG waveform deflections (e.g., P1, N1, P2, N2 etc.). However, this nomenclature is no longer recommended. That said, there is an established literature on some later ERP components such as P3a and P3b (also known as P300 or the late positive component (LPC) in the literature). In these cases, referring to their well-established names could be more appropriate (or adapted e.g., P300a, P300b), ideally citing the original article describing the component. In the auditory literature, brain-stem evoked responses were originally labelled, and today are still known, by Roman numerals I to VII.

Canonical MEEG frequency bands:

infra-slow: < 0.1 Hz
delta: 0.1 to < 4 Hz;
theta: 4 to < 8 Hz;
alpha: 8 to < 13 Hz;
beta: 13 to 30 Hz;
gamma: > 30 to 80 Hz.

Gamma band signals may occur at frequencies higher than 80 Hz[21], but the majority of MEEG studies use the lower (original) values of the range, as above. For MEG the gamma band can extend out to 1 KHz[22], so statistical analysis of gamma activity may identify *ranges of activity* within this very broad frequency band[23]. Therefore, reporting specific values of frequencies of interest within the gamma band may be more useful.

Oscillation. This term is specific to a spectral peak within a frequency band of interest, and not a general increase in MEEG power within a canonical frequency band[24]. The oscillation is defined by its peak frequency, bandwidth, and power.

Box 2. Specific MEEG terminology and definitions with respect to data analysis.

Which essential data acquisition parameters and experimental design attributes should always be reported?

When investigators report scientific findings or share data, a surprising number of important parameters are often omitted, hampering both reproducibility and replicability. To overcome these omissions, the COBIDAS MEEG report[2] contains a substantial Appendix of Tables listing desirable parameters to be reported. We do not discuss these in detail here, however; Table 1 provides a selected list of important basic descriptors of experimental paradigms, participants, and measured behaviors. We have specifically highlighted these parameters in Table because many of these tend to be omitted the most, either in already published manuscripts or in new manuscripts being submitted to journals. Here we also touch on why their omission creates ongoing problems for replications and for meta-analyses.

Issue 1: Basic hardware/software and acquisition parameters. Many published papers omit basic data acquisition details: acquisition system type, number of sensors and their spatial layout, acquisition type - continuous/epoched, sampling rate and analogue filter bandwidth (low-pass and high-pass). The latter in particular is most often omitted, yet during data acquisition *all* MEEG recording systems use filter circuitry (potentially as defaults that are not always obvious to the user) which inherently limits what is measured. Low-frequency artifacts due to respiration or skin conductance responses can be present, and on the higher frequency end, other artifacts might be aliased if they have not been filtered out (and therefore undersampled). Conversely, effects of interest in the EEG might have inadvertently been filtered out by inappropriately applied filter settings at data acquisition. There is no way to assess for these possibilities if the filter characteristics have not been reported.

Issue 2: EEG reference electrodes, impedances. A key aspect of EEG is that measurements are differential voltages relative to a reference electrode. A ground electrode serves as a way to reduce non-common mode signals in the EEG e.g. line noise or electrical stimulation artifacts. The reference and ground electrode locations must therefore always be reported.

Note that physically linked earlobe/mastoid electrodes during acquisition are not recommended as they are not a neutral reference, can introduce distortions in the data, and make modelling intractable[25]. This cannot be corrected with subsequent re-referencing or data analysis. Recording quality should also be homogenous across the scalp, and therefore the impedance measurement procedure and impedance values, for passive EEG electrode systems, should be reported. (For active electrode systems this may not always be possible). Optimal electrode impedances vary relative to an amplifier's input impedance, and to a lesser extent with electrode type (passive or active) and ambient noise level. A statement on acceptable electrode impedances (e.g. manufacturer's recommendation) for the specific setup, as well as actual values (on average, or an upper bound) and the time(s) when impedances were measured during the experiment (e.g., start, middle, end) should be provided. Reporting these procedures allows a reader to make a judgment on the quality of the data.

Issue 3: Statistical power. When null hypothesis testing is the statistical method used, reporting on *a priori* statistical power is recommended as a good practice. The probability that a study detects an effect *when there is an effect* is, however, a difficult problem in the context of EEG and MEG because it depends on the complex balance between number of trials and participants, itself a function of the experimental design (within versus between participants[26]), on chosen statistical method, and on the MEEG features of interest, including their locations, orientations and distance from sensors[27]. We recommend defining the main data feature(s) of interest and then estimating the minimal effect size to determine power. A minimal effect size is the smallest effect relevant for a given hypothesis. Effect size should be determined using estimates from independent data, existing literature, and/or pilot data. The latter should not be part of the final sample. If no electrophysiological data are available, behavioural data can be used as a minimal estimate of required sample size. In any cases, be aware that errors in calculating effect size and statistical power can occur from small sample sizes (i.e. pilot data[28]). Since (i) effect sizes of many neural effects (as measured with MEEG studies) are often smaller than that of behavioural reaction time effects, and (ii) some trials/epochs are rejected due to artifacts, thus diminishing the number of trials/epochs available for statistical analyses, this imposes lower bounds on how many trials and participants are needed[29] to achieve high statistical power. Therefore, more events and participants than has traditionally been common practice are more often required than not.

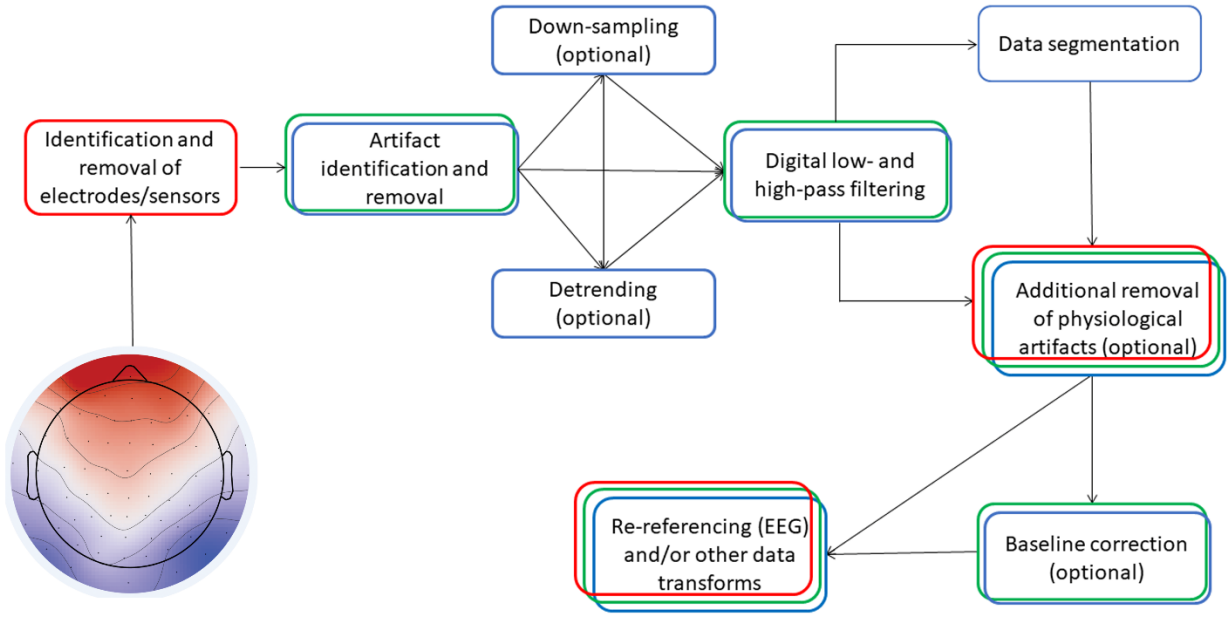
Experimental attribute	Reporting	Supplementary materials
Participant selection	- population - recruitment - sampling strategy - demographics - medications - consent	Individual demographics and questionnaires
Experimental set-up	- recording environment - seated or lying down - anaesthetic agent if any, with dosage and administration method	
Experimental task information	- Instructions - number of runs and sessions - stimuli origin and properties - software (type, version and operating system) and hardware used for stimulus presentation - conditions and stimuli order and timing - how task-relevant events are determined	scripts and stimuli
Task-free recordings	- eyes open vs closed	

	- if eyes open, fixation point or not	
Behavioural measures	- nature of the response - acquisition device (product name, model, manufacturer, recording parameters) - interface with MEEG data and calibration procedures - errors and outliers handling - statistical analyses	Individual response logs with scripts for behavioural data analysis

Table 1. Recommendations for basic experimental attributes to include in an article, along with suggested supplementary materials for increasing reproducibility.

Critical considerations for MEEG data pre-processing

We define data preprocessing as any manipulation and transformation of the data. Preprocessing order influences both the qualitative (e.g. SNR) and quantitative (e.g. deflection and spectral amplitudes) properties of the data, and thus impacts directly the replicability (Table 2). As parameter and algorithm complexity grow for MEEG data analysis, providing details about all computations is mandatory as minor changes can lead to large differences[30] in analysed output. Figure 2 outlines one of the more typical workflows, or sequence of preprocessing steps; specific recommendations for each step are available in the COBIDAS report (https://cobidasmeeg.wordpress.com/). For specific analyses, or due to specific data characteristics, the processing order can vary, but the order should be clearly justified and described in detail in accordance with our recommendations.



187
188 ***Figure 2. Standard MEEG preprocessing steps.*** *Each step affects the data in the space (red), time (blue)*
189 *and/or frequency (green) domains. Deviations from the proposed order are possible given the experimental*
190 *set-up and/or MEEG feature(s) investigated but should be justified.*

191

192

Step	Parameters	Impact
Sensor removal	- detection method and criteria - interpolation parameters if performed at this stage (e.g., trilinear, spline (+ order))	For low density coverage and/or clusters of sensors, in sensor space effects can be missed on the scalp; in source space, source locations and effects can be spurious
Artifact removal	- method used and the range of parameters (e.g., EEG data with a range larger than 75 microV) - for signal/noise separation methods (linear projection, spatial filtering techniques such as ICA[31–33]) describe the algorithm and parameters used, report the number of ICs that were obtained, how non-brain IC were identified and how back-projection was performed.	Can change or mask effects, create spurious effects
Physiological artifact removal	- types of features in the MEEG signal identified using which criteria	

	- how many (and where relative to event onset) segments were removed - MEG specific: if signal-space projection methods (SSP[34]) are used, report “empty room” measurements to estimate the topographic properties of the sensor noise and project it out from recordings containing brain activity. Related tools with a similar purpose include signal space separation methods and their temporally extended variants[35,36] that rely on the geometric separation of brain activity from noise signals in MEG data	
Downsampling	- method used (e.g. decimation, low-pass filter)	Affects the precision of time locked effect and can alter or remove spectral changes
Detrending	- detrending performed and the algorithm order (e.g., linear 1st order, piecewise, etc)	May affect connectivity metrics and statistical results
Filtering	- type of filter, cut-off frequency, filter order (or length), roll-off or transition bandwidth, passband ripple and stopband attenuation, filter delay and causality, direction of computation (one-pass forward/reverse, or two-pass forward and reverse) - for low pass, consider sampling rate setting, at least 2 to 2.5 times above the intended low pass cut off frequency (Nyquist-Shannon sampling theorem + filter roll-off)	Consequences for estimating time-courses and phases[37,38]
Segmentation	specify the length of segments	Affects connectivity values especially considering sensor vs source space[39]
Baseline correction	- assure equal baselines between conditions/groups - method used (absolute, relative, decibel, regression)	Affects signal to noise ratio, statistical type 1 error and power[40,41]

Re-referencing	- method used (subtracting the values of another channel or weighted sum of channels) - interpolation parameters if performed at this stage (e.g., trilinear, spline (+ order)) - for reference-free method (eg CSD) the software and parameter settings (interpolation method at the channel level and algorithm of the transform) must be specified.	Changes raw effect size values and statistical results
Normalization (for multivariate analyses)	- describe if performed or not - if performed, indicate the type: univariate normalization or for all channels together, i.e. multivariate normalization (or whitening). - if multivariate normalization, specify the covariance estimation procedure.	Affects source modelling and decoding performance[42,43].
Spectral transformation	- data acquisition rate must be at least twice (Nyquist theorem) the highest frequency of interest in the analyzed data - an adequate pre-stimulus baseline should be specified for evoked MEEG data i.e. the baseline duration should be equal to at least three cycles of the lowest frequency to be examined[44]. - details on the transformation algorithm and associated parameters. - the required frequency resolution is defined as the minimum frequency interval that two distinct underlying oscillatory components need to have in order to be dissociated in the analysis[45,46].	Affects the precision of results

Table 2. Overview of data preprocessing steps, parameters that should be reported and their impact on reproducibility.

Source modelling: Source modelling and reconstruction is a major processing pipeline step prior to statistical analyses and/or modeling that must be reported fully (Fig. 3). Neural source reconstruction aims at explaining the spatio-temporal pattern of observed sensor space MEEG data in terms of the underlying neuronal generators. This is known as solving the *inverse problem*, which has no unique solution (i.e. it is mathematically ill-posed). Models used to solve this problem are thus constrained by various assumptions, two important ones being the volume conduction model of the head and the source model itself. Since both affect result accuracy and reliability[47–49], details on the forward model (head model, numerical method (boundary/finite element), and conductivity), source model (distributed/ focal) and the source localization method with parameters used (e.g.,

the regularization parameter) must be reported along with the used (versioned) software for a complete and reproducible report. Information on reconstruction quality is also crucial. For both MEG and EEG, since there are multiple methods to estimate sources, the expected accuracy, errors and robustness (as described in the literature) of the chosen method should, at minimum, be described. Resampling techniques can also be used to provide further information (bias, spatial confidence intervals, etc) on the reconstruction performed with the data at hand. The source reconstruction of low-density (below 128 channels) datasets should be fully justified and interpreted with caution, given that the number of sensors impact localization accuracy[49–51] and estimation of connectivity[52]. Different source modelling methods can be advantageous for particular applications, so reporting the rationale for choosing a source model is also important.

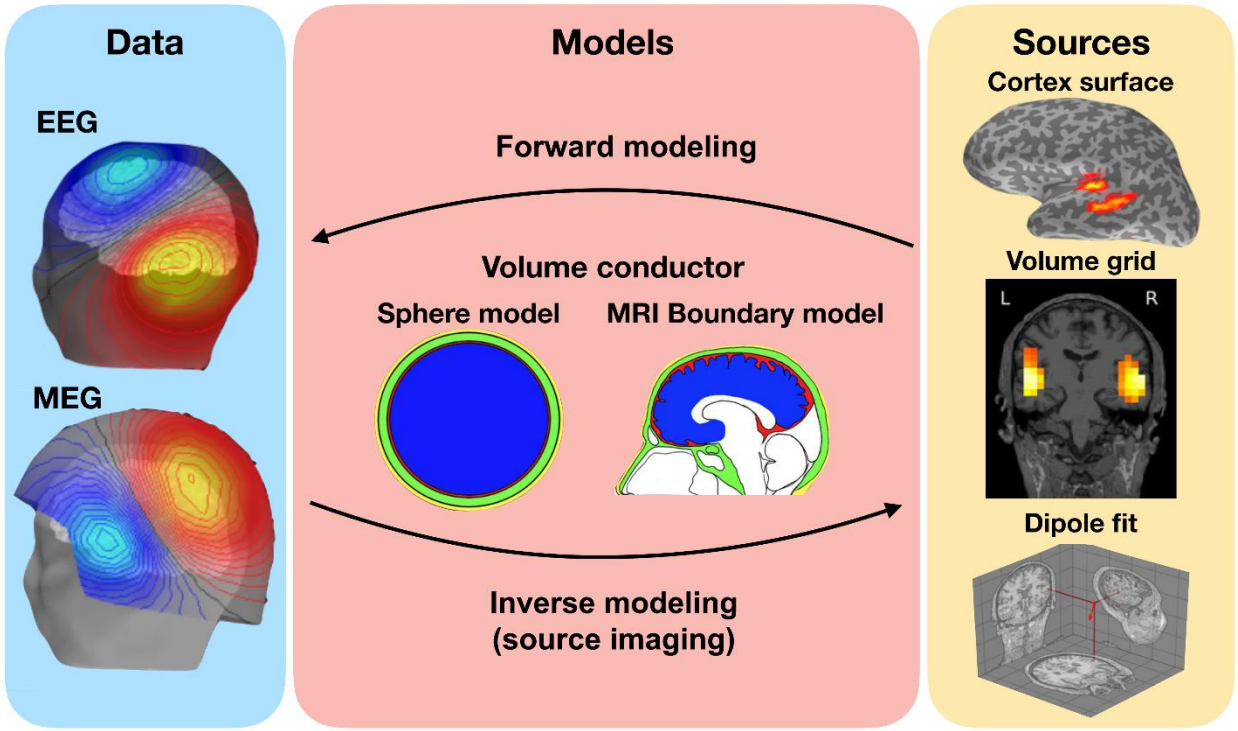


Figure 3. Illustration of source modelling approaches. *To find active neural sources, a forward model must first be used to determine the scalp distribution of the EEG potential or MEG magnetic field for a (set of) known source(s). These models vary according to how sources are defined (either on the cortical surface or on a volumetric grid) and the volume conduction model, which simulates effects on the tissues in the head on propagation of activity to MEEG sensors (spherical head model vs. MRI derived models - here showing bone (green), cerebrospinal fluid (red), gray and white matter (blue) tissues). Information from the forward model is then inverted to attribute active sources to the measured MEEG signals.*

Critical considerations for MEEG data processing

We define data processing as mathematical procedures that *do not change* the data, i.e. statistical analysis and statistical modeling. There are many valid methods to analyse MEEG data. The chosen method should best answer the posed scientific question[53] and a rationale for its use should always be provided. Here we briefly examine some of the main data processing issues discussed in the COBIDAS MEEG report[2].

ROI-based analyses: Selecting specific channels or source-level Regions-Of-Interest (ROI) based on grand average differences between conditions/groups and then performing statistical tests on these has at times been seen in the MEEG literature. This, however, creates estimation biases (i.e. "double-dipping")[54,55], irrespective of whether one works in sensor or source space. ROI analyses in time, frequency or space (peak analysis, window average, etc) while legitimate, should be justified *a priori* based on prior literature or independent data or statistical contrasts.

Mass univariate statistical modelling: More recently, analyses tend to be performed at the participant and group levels, using a hierarchical or mixed model approach for the whole data volume (3D source space), and/or the spatio-temporal sensor space[56,57]. These types of analyses (and those that follow in the subsequent sections below) have become more common and have not typically been addressed in previous guidelines. Compared to tomographic methods, MEEG can have missing data (e.g., bad channels, or transient intervals with artifacts), so reporting on how missing data have been treated is crucial. Results must be corrected for multiple testing/comparisons (e.g., full brain analyses or multiple feature/component maxima), but both *a priori* and *a posteriori* thresholds[58] cannot adequately control the Type 1 family-wise error and should be avoided[59]. Special attention must also be given to data smoothness when using random field theory[60]. This is in contrast to *a posteriori* thresholds using null distributions (bootstrap and permutations), which control well for family-wise Type 1 error rate[61,62].

Multivariate statistical inference: Multivariate statistical tests (e.g. MANOVA, Linear Discriminant Analysis) are typically performed in space or time/frequency, thus also leading to a multiple comparisons problem that needs to be properly addressed. The problem of not correcting adequately for multiple comparisons remains a common omission for such data analyses.

Multivariate pattern classification: Decoding approaches should strive to minimise bias and unrealistically high classification rates, commonly referred to as "overfitting". To avoid overfitting, a nested cross-validation procedure should be used, where independent subsets of the data are used to estimate the parameters, fit the classification model, and estimate performance metrics. It is also important to justify data-split choice, as some approaches can give biased estimates (e.g. leave-one-out on correlated data[63]).

Connectivity: The term "connectivity" is an umbrella term often used to refer to multiple methods, which may create some confusion in the literature[64,65]. In the MEEG context, it generally refers to analyses that aim to detect *coupling* between two or more channels or sources. We recommend explicitly referring to functional (correlational) or effective

(causal) connectivity[66] and to describe the specific method used (e.g. effective Granger connectivity, partial coherence, dynamic causal modelling (DCM), etc). Table 3 outlines different approaches in connectivity analyses and lists important variables to report. With respect to the *computed metrics*[67]*,* it is essential to report all parameters since they have a major effect on analytic outputs[49,52]. Statistical dependence measures in either sensor or source space should be specified (e.g., correlation, phase coupling, amplitude coupling, spectral coherence, entropy, DCM, Granger causality), as well as analysis assumptions (e.g., linear versus unspecified; directional versus non-directional). For cross-frequency coupling (CFC)-based analyses, coupling type[68] should be explicitly noted. CFC occurs when activity at lower frequencies modulates higher frequency amplitude, phase or frequency. Since even one type of CFC can be extracted using multiple methods[69–71], analysis methods and all associated parameters, such as filtering, must also be specified in detail.

Connectivity from MEG or EEG can be obtained from *sensor or source space measures*[72]*,* and many discussions on the validity or utility of those measures exist. Our view is that while statistical metrics of dependency can be calculated at channel level (which can be useful for e.g. biomarking), these are not measures of neural connectivity[67,73] and therefore cannot be used for causal inference[74]. Neural connectivity can only be obtained after biophysical modeling (assuming it is accurate enough), considering volume conduction (e.g. spatial leakage of source signals[76]) and spurious connections due to unobserved common sources.

Specifications	Parameters
Connectivity analysis	specify type: effective [causal] or functional [correlational] specify exact method used
Network estimation approaches	approach: data driven [e.g. ICA, time frequency analysis based] or anatomical/model driven? native space vs. template space?[75,76] If data driven, specify methods & parameters [e.g. time-frequency decomposition method] if anatomically driven, specify parcellation approach & parameters graph theoretical measures: motivation of metrics[77], specify if directed/undirected network, define nodes/edges, specify thresholding criteria
Consideration on computing metrics	consider effects of epoch length[39] for dynamic connectivity measures describe all temporal parameters[78] (e.g. window size, overlap, wavelet frequency and scale) for spectral coherence/synchrony measures: specify exact formulation (or reference), any subtraction or normalisation with respect to an experimental condition or mathematical criterion, is the measure debiased? for partial coherence and multiple coherence measures: describe all variables, specify exact variables used, and whether data are partialised, marginalised, conditioned, or orthogonalized for DCM[79] specify model type (event-related potential, canonical

	microcircuit); describe full space of considered functional architectures; connectivity matrices present/modulated (forward, backward, lateral, if intrinsic); vector of between-trial effects, the number of modes, the temporal window modelled, and the priors on source locations; statistical approach: at the level of models or the family of models (Fixed- or Random-effects, FFX or RFX); connectivity parameters (Frequentist versus Bayesian, Bayesian Model Averaging (BMA) over all models or conditioned on the winning family/model

Table 3. Necessary parameters to report in MEEG connectivity modeling to ensure reproduction of the method used.

Results reporting and display items

The COBIDAS MEEG report[2] discusses results reporting and figures in considerable detail. In what follows we highlight some of the more common problematic aspects, where even previously published neurophysiological studies have omitted important data characteristics.

Issue 1: Figures. In figures depicting neurophysiological waveforms, we advocate the inclusion of variability measures (e.g., confidence intervals) and clearly annotated scales for all displayed data attributes. Moreover, since MEEG activity is characterized by its topography, it is recommended that waveforms/spectra of the full set of channels are shown (either in the main document or in supplementary materials).

Issue 2: Using frequency band names across the lifespan. Considerable ambiguities and confusion exist in the spontaneous/resting-state MEEG literature due to inconsistent use of terminology, and not assessing a particular cortical rhythm's reactivity[16]. The well-known posterior alpha rhythm characteristically occurs following eye closure and diminishes greatly on eye opening. Importantly, during the lifespan posterior alpha changes peak frequency: in infants (3-4 months of age) a reactive posterior rhythm first appears at ~4 Hz, increasing to ~6 Hz at 12 months of age and to ~8 Hz at 36 months, reaching adult frequencies of ~10 Hz by 6-12 years[80], and slowing again with normal ageing[20]. Specifying the *frequency and distribution of the activity* and noting its reactivity is therefore important when studying aging. To reduce confusion, terms such as "baby alpha" should be avoided, as central/rolandic ("mu") rhythms (see COBIDAS MEEG report for other issues related to mu rhythms) can develop in infants *before* the posterior reactive rhythm that ultimately becomes fully-fledged "alpha" is seen. Currently, it is difficult to perform meta-analyses because of the variability of use of various frequency band names in the literature.

Issue 3: Underspecifying results of statistical analyses. For group or experimental condition differences, the test statistic (e.g., F-values, t-values, Bayes Factors) must be displayed. Reporting model assumptions (e.g. in linear models this includes Gaussianity of residuals) and effect size (e.g., Cohen's d, percentage difference and/or raw magnitude) are also encouraged. It is also good practice to report the explained model variance and data fit (both R-squared and RMSE), as well as parameters deriving from the model(s) (e.g., weight estimates, maximum statistical values). For predictive models,

decoding accuracy (classification), R-squared or RMSE (regression) are the measures of choice, but chance level should be included[81]. The area under a ROC curve can also be used when doing binary classification. Whichever method is used, each (expected) effect should be reported, *significant or not*, allowing readers to evaluate the dataset. This permits comparison with similar studies, facilitates informed power analyses for planning future studies, and will enable developments of a quantitative, more reproducible, view of brain dynamics[82].

For mass-univariate and multivariate analyses, statistical maps of the space tested are usually displayed, with corresponding waveforms and topographic maps. While statistical significance matters, providing only thresholded maps limits reproducibility. We recommend displaying thresholded maps in manuscripts (with description of thresholding method), while providing raw maps for all channels and time/frequency frames in supplementary materials (ideally as a data matrix in a repository and not just a figure). To allow the reader to evaluate observed effects, both the time course of the model parameters and underlying data should be made available. Consideration should be given to what figures should appear in the main manuscript versus those appearing in the Supplementary Materials section.

The evolution of COBIDAS, data sharing and future neuroimaging studies

The current COBIDAS MEEG recommendations correspond to best practices in 2019. Reporting data using these criteria should improve the generation of reproducible and replicable findings. As MEEG analysis pipelines become increasingly more complex, more methodological details will likely need to be reported, challenging current views on good writing practice and journal policies. In anticipation of, and to facilitate, this process COBIDAS MEEG is a 'living' document (https://cobidasmeeg.wordpress.com/), that will have periodic updates to include best practices for new methods as they become more established.

We also encourage the MEEG community to share raw and derived data using BIDS, together with data processing scripts[83]. Sharing of data and scripts fosters reproducibility and script re-usage encourages replicability across laboratories, promoting benefits to research training and education. A huge challenge to MEEG replicability is the large data space and variety of methods. Sharing of derived MEEG data (similar to fMRI data where statistical maps are shared) would allow direct comparisons, replications and aggregations of results across studies (e.g., meta-analysis). In an era of electronic publishing, sharing derived data is straightforward (e.g. grand average ERPs between two conditions consist of a file of a few kilobytes that can be added as supplementary material or posted in a data repository).

Sharing original data is not always feasible since participant consent is required and issues of confidentiality may be a particular concern for clinical samples. Datasets with *whole head* anatomical MRI data can be similarly problematic, as head models cannot be reconstructed if T1-weighted images are defaced or skull stripped. Even without structural

MRI, functional imaging data, including MEEG[84], could be indirectly identifiable. Confidentiality is currently a world-wide discussion point, with cross-continental data-sharing initiatives posing some challenges. We strongly encourage seeking ethical clearance from participants regarding data sharing *before* commencing any study (see open brain consent form examples (https://open-brain-consent.readthedocs.io/) for easy to follow templates).

Exciting technical developments in MEEG (Fig. 1) will require updating of the COBIDAS report to include best, modern, practices for these new methods, in particular for machine learning algorithms that will likely play an increasingly prominent role in years to come[85,86]. Similarly, new generation room temperature MEG measurement sensors (or optically pumped magnetometers) are emerging, allowing previously unavailable flexible configurations of MEG sensor arrays[87,88]. As we also progress towards "putting the brain back into the body", multimodal integration of MEEG data with other technologies such as the simultaneous recording of movements or autonomic nervous responses, will create new challenges in best practices, as cognitive and systems neuroscience moves out of the laboratory, to more ecologically valid scenarios, and "into the wild".

Conclusions

The first COBIDAS MEEG report was completed with prolonged and extensive collaboration and consultation within the neuroimaging community. We aimed to compile best practices for data gathering, analysis and sharing, to improve scientific reproducibility and replicability. These guidelines were constructed not only for preparation of manuscripts and grants, but also for scientists serving in editing and review roles, as well as for education and research training of future scientists. Like the COBIDAS MRI report, we see the COBIDAS MEEG report as a living document - designed to keep pace with ever-changing scientific and methodological developments in the field. OHBM will continue its efforts in defining best practices for brain imaging and welcomes all to participate and contribute to this endeavour.

Authorship

CP and AP chaired the committee, planned the overall structure of the COBIDAS document and this manuscript, each author contributed to entire sections of the COBIDAS document used for this manuscript, all authors contributed and reviewed this manuscript.

Acknowledgement

The Committee thanks the hundreds of OHBM members who provided feedback on the early version of the report and on the website. Thank you to Tom Nichols, Chair, OHBM COBIDAS MRI for his insightful comments on an earlier draft of this Perspective.

References

1. Barba, L. A. Terminologies for Reproducible Research. *arXiv:1802.03311 [cs]* (2018).
2. Pernet, C. R. *et al. Best Practices in Data Analysis and Sharing in Neuroimaging using MEEG*. https://osf.io/a8dhx (2018) doi:10.31219/osf.io/a8dhx.
3. Nichols, T. E. *et al.* Best practices in data analysis and sharing in neuroimaging using MRI. *Nature Neuroscience* **20**, 299–303 (2017).
4. Gorgolewski, K. J. *et al.* The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data* **3**, 160044 (2016).
5. Niso, G. *et al.* MEG-BIDS, the brain imaging data structure extended to magnetoencephalography. *Scientific Data* **5**, 180110 (2018).
6. Pernet, C. R. *et al.* EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Scientific Data* **6**, 103 (2019).
7. Holdgraf, C. *et al.* iEEG-BIDS, extending the Brain Imaging Data Structure specification to human intracranial electrophysiology. *Scientific Data* **6**, 102 (2019).
8. Donchin, M. *et al.* Publication criteria for studies of evoked potentials (EP) in man: Methodology and publication criteria. in *Progress in Clinical Neurophysiology: Attention, Voluntary Contraction and Event-Related Cerebral Potentials.* (ed. Desmedt, J. E.) vol. 1 1–11 (1977).
9. Pivik, R. T. *et al.* Guidelines for the recording and quantitative analysis of electroencephalographic activity in research contexts. *Psychophysiology* **30**, 547–558 (1993).
10. Picton, T. W. *et al.* Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology* **37**, 127–152 (2000).
11. Duncan, C. C. *et al.* Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology* **120**, 1883–1908 (2009).
12. Gross, J. *et al.* Good practice for conducting and reporting MEG research. *NeuroImage* **65**, 349–63 (2013).
13. Keil, A. *et al.* Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology* **51**, 1–21 (2014).
14. Kane, N. *et al.* A revised glossary of terms most commonly used by clinical electroencephalographers and updated proposal for the report format of the EEG findings. Revision 2017. *Clinical Neurophysiology Practice* **2**, 170–185 (2017).
15. Hari, R. *et al.* IFCN-endorsed practical guidelines for clinical magnetoencephalography (MEG). *Clinical Neurophysiology* **129**, 1720–1747 (2018).
16. Hari, R. & Puce,A. *MEG-EEG Primer*. (Oxford University Press, 2017).
17. Jobert, M. *et al.* Guidelines for the Recording and Evaluation of Pharmaco-EEG Data in Man: The International Pharmaco-EEG Society (IPEG). *NPS* **66**, 201–220 (2012).
18. Berger, H. Über das Elektrenkephalogramm des Menschen. *Archiv f. Psychiatrie* **87**, 527–570 (1929).
19. Walter, W.G. Ther location od cerebral tumors by electroencephalography. *Lancet* 305–308 (1936).
20. Krishnan, V., Chang, B. S. & Schomer, D. L. Normal EEG in Wakefulness and Sleep: Adults and Elderly. in *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields* (eds. Schomer, D.L. & Lopes da Silva, F.H.) (Oxford University Press).
21. Amzica, F. & Silva, F. H. L. da. Cellular Substrates of Brain Rhythms. in *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields* (eds. Schomer, D.L. & Silva,. F) chapter 2 (Oxford University Press, 2018).

22. Baillet, S. Magnetoencephalography for brain electrophysiology and imaging. *Nat Neurosci* **20**, 327–339 (2017).
23. Uhlhaas, P. J., Pipa, G., Neuenschwander, S., Wibral, M. & Singer, W. A new look at gamma? High- (>60 Hz) γ-band activity in cortical networks: Function, mechanisms and impairment. *Progress in Biophysics and Molecular Biology* **105**, 14–28 (2011).
24. Lopes da Silva, F. EEG and MEG: Relevance to Neuroscience. *Neuron* **80**, 1112–1128 (2013).
25. Katznelson, R.D. EEG recording, electrode placement, and aspects of generator localization. in *Electric Fields of the Brain. The Neurophysics of EEG* (ed. Nunez, P.) 176–213 (Oxford University Press, 1981).
26. Boudewyn, M. A., Luck, S. J., Farrens, J. L. & Kappenman, E. S. How many trials does it take to get a significant ERP effect? It depends. *Psychophysiology* **55**, e13049 (2018).
27. Chaumon, M., Puce, A. & George, N. Statistical power: implications for planning MEG studies. *bioRxiv* 852202 (2020) doi:10.1101/852202.
28. Albers, C. & Lakens, D. When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. *Journal of Experimental Social Psychology* **74**, 187–195 (2018).
29. Brysbaert, M. & Stevens, M. Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition* **1**, 9 (2018).
30. Robbins, K. A., Touryan, J., Mullen, T., Kothe, C. & Bigdely-Shamlo, N. How Sensitive Are EEG Results to Preprocessing Methods: A Benchmarking Study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **28**, 1081–1090 (2020).
31. Brown, G. D., Yamada, S. & Sejnowski, T. J. Independent component analysis at the neural cocktail party. *Trends in Neurosciences* **24**, 54–63 (2001).
32. Jung, T. P. *et al.* Imaging brain dynamics using independent component analysis. *Proceedings of the IEEE* **89**, 1107–1122 (2001).
33. Onton, J., Westerfield, M., Townsend, J. & Makeig, S. Imaging human EEG dynamics using independent component analysis. *Neuroscience & Biobehavioral Reviews* **30**, 808–822 (2006).
34. Uusitalo, M. A. & Ilmoniemi, R. J. Signal-space projection method for separating MEG or EEG into components. *Med. Biol. Eng. Comput.* **35**, 135–140 (1997).
35. Taulu, S., Kajola, M. & Simola, J. Suppression of Interference and Artifacts by the Signal Space Separation Method. *Brain Topogr* **16**, 269–275 (2004).
36. Taulu, S. & Simola, J. Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. *Phys. Med. Biol.* **51**, 1759–1768 (2006).
37. Rousselet, G. A. Does Filtering Preclude Us from Studying ERP Time-Courses? *Front. Psychol.* **3**, (2012).
38. Widmann, A., Schröger, E. & Maess, B. Digital filter design for electrophysiological data – a practical approach. *Journal of Neuroscience Methods* **250**, 34–46 (2015).
39. Fraschini, M. *et al.* The effect of epoch length on estimated EEG functional connectivity and brain network organisation. *J. Neural Eng.* **13**, 036015 (2016).
40. Grandchamp, R. & Delorme, A. Single-Trial Normalization for Event-Related Spectral Decomposition Reduces Sensitivity to Noisy Trials. *Front. Psychol.* **2**, (2011).
41. Alday, P. M. How much baseline correction do we need in ERP research? Extended GLM model can replace baseline correction while lifting its limits. *Psychophysiology* **56**, e13451 (2019).
42. Engemann, D. A. & Gramfort, A. Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals. *Neuroimage* **108**, 328–342 (2015).
43. Guggenmos, M., Sterzer, P. & Cichy, R. M. Multivariate pattern analysis for MEG: A comparison of dissimilarity measures. *NeuroImage* **173**, 434–447 (2018).
44. Cohen, M. *Analyzing Neural Time Series Data. Theory and Practice.* (MIT Press, 2014).

45. Bloomfield, P. *Fourier Analysis of Time Series: An Introduction*. (Wiley, 2013).
46. Boashash, B. *Time-frequency Signal Analysis and Processing: A Comprehensive Reference*. (Elsevier, 2003).
47. Baillet, S., Mosher, J. C. & Leahy, R. M. Electromagnetic brain mapping. *IEEE Signal Processing Magazine* **18**, 14–30 (2001).
48. Michel, C. & He, B. EEG Mapping and Source Imaging. in *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields* (eds. Schomer, D. L. & Silva, F. H. L. da) chap 45 (Oxford University Press, 2018).
49. Michel, C. M. *et al.* EEG source imaging. *Clinical Neurophysiology* **115**, 2195–2222 (2004).
50. Michel, C. M. & Brunet, D. EEG Source Imaging: A Practical Review of the Analysis Steps. *Front. Neurol.* **10**, (2019).
51. Brodbeck, V. *et al.* Electroencephalographic source imaging: a prospective study of 152 operated epileptic patients. *Brain* **134**, 2887–2897 (2011).
52. Hassan, M., Dufor, O., Merlet, I., Berrou, C. & Wendling, F. EEG Source Connectivity Analysis: From Dense Array Recordings to Brain Networks. *PLOS ONE* **9**, e105041 (2014).
53. Kass, R. E. *et al.* Ten Simple Rules for Effective Statistical Practice. *PLOS Computational Biology* **12**, e1004961 (2016).
54. Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F. & Baker, C. I. Circular analysis in systems neuroscience: the danger of double dipping. *Nature Neuroscience* **12**, 535–540 (2009).
55. Kriegeskorte, N., Lindquist, M. A., Nichols, T. E., Poldrack, R. A. & Vul, E. Everything You Never Wanted to Know about Circular Analysis, but Were Afraid to Ask. *J Cereb Blood Flow Metab* **30**, 1551–1557 (2010).
56. Kilner, J. M., Kiebel, S. J. & Friston, K. J. Applications of random field theory to electrophysiology. *Neuroscience letters* **374**, 174–178 (2005).
57. Pernet, C. R., Chauveau, N., Gaspar, C. & Rousselet, G. A. LIMO EEG: A Toolbox for Hierarchical LInear MOdeling of ElectroEncephaloGraphic Data. *Computational Intelligence and Neuroscience* **2011**, 1–11 (2011).
58. Guthrie, D. & Buchwald, J. S. Significance Testing of Difference Potentials. *Psychophysiology* **28**, 240–244 (1991).
59. Piai, V., Dahlslätt, K. & Maris, E. Statistically comparing EEG/MEG waveforms through successive significant univariate tests: How bad can it be? *Psychophysiology* **52**, 440–443 (2015).
60. Eklund, A., Nichols, T. E. & Knutsson, H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences* **113**, 7900–7905 (2016).
61. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods* **164**, 177–190 (2007).
62. Pernet, C. R., Latinus, M., Nichols, T. E. & Rousselet, G. A. Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study. *Journal of Neuroscience Methods* **250**, 85–93 (2015).
63. Varoquaux, G. *et al.* Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. *NeuroImage* **145**, 166–179 (2017).
64. O'Neill, G. C. *et al.* Dynamics of large-scale electrophysiological networks: A technical review. *NeuroImage* **180**, 559–576 (2018).
65. He, B. *et al.* Electrophysiological Brain Connectivity: Theory and Implementation. *IEEE Trans Biomed Eng* (2019) doi:10.1109/TBME.2019.2913928.
66. Friston, K. J. Functional and effective connectivity: A review. *Brain Connectivity* **1**, 13–36 (2011).
67. Haufe, S., Nikulin, V. V., Müller, K.-R. & Nolte, G. A critical assessment of connectivity measures for EEG data: A simulation study. *NeuroImage* **64**, 120–133 (2013).

68. Jensen, O. & Colgin, L. L. Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences* **11**, 267–269 (2007).
69. Tort, A.B.L., Komorowski, R., Eichenbaum, H. & Kopell, N.J. Measuring Phase-Amplitude Coupling Between Neuronal Oscillations of Different Frequencies | Journal of Neurophysiology. *Journal of Neurophysiology* **104**, 1195–1210 (2010).
70. van Wijk, B. C. M., Jha, A., Penny, W. & Litvak, V. Parametric estimation of cross-frequency coupling. *Journal of Neuroscience Methods* **243**, 94–102 (2015).
71. Dupré la Tour, T. *et al.* Non-linear auto-regressive models for cross-frequency coupling in neural time series. *PLOS Computational Biology* **13**, e1005893 (2017).
72. Lai, M., Demuru, M., Hillebrand, A. & Fraschini, M. A comparison between scalp- and source-reconstructed EEG networks. *Sci Rep* **8**, 1–8 (2018).
73. Valdes-Sosa, P. A., Roebroeck, A., Daunizeau, J. & Friston, K. Effective connectivity: Influence, causality and biophysical modeling. *NeuroImage* **58**, 339–361 (2011).
74. Reid, A. T. *et al.* Advancing functional connectivity research from association to causation. *Nature Neuroscience* **22**, 1751–1760 (2019).
75. Farahibozorg, S.-R., Henson, R. N. & Hauk, O. Adaptive cortical parcellations for source reconstructed EEG/MEG connectomes. *Neuroimage* **169**, 23–45 (2018).
76. Mahjoory, K. *et al.* Consistency of EEG source localization and connectivity estimates. *NeuroImage* **152**, 590–601 (2017).
77. Sporns, O. Contributions and challenges for network models in cognitive neuroscience. *Nat Neurosci* **17**, 652–660 (2014).
78. Tewarie, P. *et al.* Tracking dynamic brain networks using high temporal resolution MEG measures of functional connectivity. *NeuroImage* **200**, 38–50 (2019).
79. Litvak, V. *et al.* EEG and MEG Data Analysis in SPM8. *Comput Intell Neurosci* **2011**, (2011).
80. Pearl, P. L. *et al.* Normal EEG in Wakefulness and Sleep: Preterm; Term; Infant; Adolescent. in *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields* (eds. Schomer, D.L. & Lopes da Silva, F.H.) (Oxford University Press).
81. Jas, M. *et al.* A Reproducible MEG/EEG Group Study With the MNE Software: Recommendations, Quality Assessments, and Good Practices. *Front. Neurosci.* **12**, (2018).
82. Rousselet, G. & Pernet, C. Quantifying the Time Course of Visual Object Processing Using ERPs: It's Time to Up the Game. *Frontiers in Psychology* **2**, 107 (2011).
83. Eglen, S. J. *et al.* Toward standard practices for sharing computer code and programs in neuroscience. *Nat Neurosci* **20**, 770–773 (2017).
84. Leppäaho, E. *et al.* Discovering heritable modes of MEG spectral power. *Human Brain Mapping* **40**, 1391–1402 (2019).
85. Tuckute, G., Hansen, S. T., Pedersen, N., Steenstrup, D. & Hansen, L. K. Single-Trial Decoding of Scalp EEG under Natural Conditions. *Computational Intelligence and Neuroscience* https://new.hindawi.com/journals/cin/2019/9210785/ (2019) doi:https://doi.org/10.1155/2019/9210785.
86. Pion-Tonachini, L., Kreutz-Delgado, K. & Makeig, S. The ICLabel dataset of electroencephalographic (EEG) independent component (IC) features. *Data in Brief* **25**, 104101 (2019).
87. Boto, E. *et al.* A new generation of magnetoencephalography: Room temperature measurements using optically-pumped magnetometers. *NeuroImage* **149**, 404–414 (2017).
88. Boto, E. *et al.* Moving magnetoencephalography towards real-world applications with a wearable system. *Nature* **555**, 657–661 (2018).