Neuroscience and Biobehavioral Reviews 84 (2018) 151–161

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

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

Review article

Ten simple rules for neuroimaging meta-analysis

Veronika I. Müller[a,b,*,1], Edna C. Cieslik[a,b,*,1], Angela R. Laird[c], Peter T. Fox[d,e,f], Joaquim Radua[g,h,i], David Mataix-Cols[h], Christopher R. Tench[j], Tal Yarkoni[k], Thomas E. Nichols[l,m], Peter E. Turkeltaub[n,o], Tor D. Wager[p,q], Simon B. Eickhoff[a,b]

[a] Institute of Systems Neuroscience and Institute of Clinical Neuroscience & Medical Psychology, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[b] Institute of Neuroscience und Medicine (INM-1, INM-7), Research Centre Jülich, Jülich, Germany
[c] Department of Physics, Florida International University, Miami, FL, USA
[d] Research Imaging Institute, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA
[e] Research Service, South Texas Veterans Administration Medical Center, San Antonio, TX, USA
[f] Shenzhen Institute of Neuroscience, Shenzhen University, Shenzhen, China
[g] FIDMAG Germanes Hospitalàries, CIBERSAM, Barcelona, Spain
[h] Department of Clinical Neuroscience, Centre for Psychiatry Research, Karolinska Institutet, Stockholm, Sweden
[i] Department of Psychosis Studies, Institute of Psychology, Psychiatry, and Neuroscience, King's College London, London, United Kingdom
[j] (CRT) Division of Clinical Neurosciences, Clinical Neurology, University of Nottingham, Queen's Medical Centre, Nottingham, United Kingdom
[k] Department of Psychology, University of Texas at Austin, Austin, TX, USA
[l] Department of Statistics, University of Warwick, Coventry, United Kingdom
[m] Warwick Manufactoring Group, University of Warwick, Coventry, United Kingdom
[n] Department of Neurology, Georgetown University Medical Center, Washington, DC, USA
[o] Research Division, MedStar National Rehabilitation Hospital, Washington, DC, USA
[p] Department of Psychology and Neuroscience, University of Colorado, Boulder, Colorado, USA
[q] Institute of Cognitive Science, University of Colorado, Boulder, USA

ARTICLE INFO

Keywords:
Meta-analysis
Ten simple rules
Guidelines
Neuroimaging
fMRI
PET

ABSTRACT

Neuroimaging has evolved into a widely used method to investigate the functional neuroanatomy, brain-behaviour relationships, and pathophysiology of brain disorders, yielding a literature of more than 30,000 papers. With such an explosion of data, it is increasingly difficult to sift through the literature and distinguish spurious from replicable findings. Furthermore, due to the large number of studies, it is challenging to keep track of the wealth of findings. A variety of meta-analytical methods (coordinate-based and image-based) have been developed to help summarise and integrate the vast amount of data arising from neuroimaging studies. However, the field lacks specific guidelines for the conduct of such meta-analyses. Based on our combined experience, we propose best-practice recommendations that researchers from multiple disciplines may find helpful. In addition, we provide specific guidelines and a checklist that will hopefully improve the transparency, traceability, replicability and reporting of meta-analytical results of neuroimaging data.

1. Introduction

Over the last two decades, neuroimaging has evolved into a widely used method to investigate functional neuroanatomy, brain-behaviour relationships, and pathophysiology of brain disorders. However, single imaging studies usually rely on underpowered studies with small sample sizes, which leads to many missed results (Button et al., 2013) and pushes researchers towards analyses and thresholding procedures that increase false positives (Eklund et al., 2016; Wager et al., 2007; Wager et al., 2009; Woo et al., 2014). In addition, results are strongly influenced by experimental and analyses procedures (Carp, 2012) and replication studies are rare. Thus, it is increasingly difficult to sift through the enormous neuroimaging literature and distinguish spurious from replicable findings, and even harder to gauge whether effects in individual studies can be generalized to a task or patient group in a way that is robust to variation in the specific task and details of analysis choices performed. Furthermore, due to the large number of studies, it is challenging to keep track of the wealth of findings (Radua and

* Corresponding author at: Institute of Neuroscience and Medicine (INM-1, INM-7), Research Centre Jülich, Wilhelm-Johnen-Straße, D-52428 Jülich, Germany.
E-mail addresses: v.mueller@fz-juelich.de (V.I. Müller), e.cieslik@fz-juelich.de (E.C. Cieslik).
[1] Authors contributed equally to this work.

https://doi.org/10.1016/j.neubiorev.2017.11.012
Received 24 April 2017; Received in revised form 18 November 2017; Accepted 20 November 2017
Available online 24 November 2017

Mataix-Cols, 2012). Thus, there is a need to quantitatively consolidate effects across individual studies in order to overcome problems associated with individual neuroimaging studies.

One potent approach to synthesizing the multitude of results in an unbiased fashion is to perform a meta-analysis. There are two general approaches to neuroimaging meta-analyses: image-based and coordinate-based meta-analyses. Image-based meta-analyses are based on the full statistical images of the original studies, whereas coordinate-based meta-analyses only use the x,y,z-coordinates (and in some cases their z-statistic) of each peak location reported in the respective publication. Image-based meta-analyses allow for the use of hierarchical mixed effects models that account for intra-study variance and random inter-study variation (Salimi-Khorshidi et al., 2009) as the full information required for this is provided in image form. However, due to the fact that whole-brain statistical images are rarely shared (but see Gorgolewski et al., 2015; http://neurovault.org, for recent approaches of sharing unthresholded statistical images in an online database), most meta-analytic research questions cannot yet be addressed with image-based meta-analysis. In contrast, while coordinate-based meta-analyses use a sparser representation of findings, almost all individual neuroimaging studies provide their results as coordinates in standardized anatomical space (either MNI (Collins et al., 1994) or Talairach (Talairach and Tournoux, 1988) space). Thus, coordinate-based meta-analyses allow us to capitalize on much of the published neuroimaging literature, and provide a quantitative summary of these results to answer a specific research question.

There are different approaches to coordinate based meta-analysis, including (multilevel) kernel density analysis (KDA, MKDA; e.g., Wager et al., 2004; Wager et al., 2007; Pauli et al., 2016), gaussian-process regression (GPR; Salimi-Khorshidi et al., 2011), activation likelihood estimation (ALE; Eickhoff et al., 2012; Eickhoff et al., 2009; Turkeltaub et al., 2002; Turkeltaub et al., 2012), parametric voxel-based meta-analysis (PVM; Costafreda et al., 2009), signed differential mapping (SDM; Radua and Mataix-Cols, 2009). A revised version of SDM, termed effect-size SDM (ES-SDM), also allows for the combination of coordinate-based results and statistical images (Radua et al., 2012).

Despite the increasing use of meta-analytic approaches in the last few years, there is a lack of concrete recommendations regarding how to perform neuroimaging-based meta-analyses, report findings, or make results available for the whole neuroimaging community to foster reproducibility of neuroimaging meta-analytic results. For individual MRI experiments, such guidelines have already been developed (COBIDAS; Nichols et al., 2017). However, best practices for neuroimaging meta-analyses differ from those of individual imaging studies (and also from those of effect-size based meta-analyses of behavioral studies, (e.g., MARS; (American Psychological Association, 2010))). Thus, the aim of this paper is twofold. First, we provide best-practice recommendations that should be considered carefully when performing neuroimaging meta-analyses and help researchers to make informed and traceable decisions. Second, we set standards regarding which information should be reported when publishing meta-analyses to enable other researchers to replicate the study. While these recommendations are primarily relevant to coordinate-based meta-analyses, most of them also hold true for image-based meta-analyses.

2. Recommendations

2.1. Be specific about your research question

The critical first step of any meta-analysis is to specify as precisely as possible the research question and the approach towards investigating it. For most functional neuroimaging meta-analyses (this decision is not relevant for structural imaging studies), the researcher must first decide which paradigms to include in the meta-analysis. For example, a researcher interested in cognitive action control may want to know which regions are consistently found activated or deactivated across experiments that required participants to inhibit a prepotent response in favor of a non-routine one. For this example, the question arises if one should include all experiments that test cognitive action control, no matter what paradigm was used (e.g., Stop-signal, Go/No-Go, Stroop, Flanker tasks…), or limit the analysis to a specific paradigm (e.g., Stop-signal task). Considering the consequences for interpretation, the latter case would be specific to the cancelling of an already initiated action, while a meta-analysis across all paradigms would focus on the higher order supervisory control processes necessary in all paradigm types. Importantly, if one decides to include different paradigms, it may be helpful to ensure that the distribution of experiments is relatively balanced across tasks. However, in this context, it should be noted, that if there is enough literature available, there is the possibility to not only calculate one main meta-analysis, but rather also sub-analyses which may focus on more specialized processes (e.g., different paradigm classes) or groups (e.g. different patient samples). For example, one could plan to calculate a general meta-analysis across Stop-signal, Go/No-Go, Stroop and Flanker tasks and then also individual sub-analyses for each paradigm. Convergence across paradigms could be then tested by overlapping the results of the different sub-analyses, or quantitatively using an omnibus test of difference in reported activation pattern (Tench et al., 2014). However, these choices of sub-analyses should have a rationale and be made beforehand and not after inspecting the data (see below). Importantly, brain processes may not always be organized by named task type and minor variations in paradigms can produce large changes in cognitive strategies. As an example, Gilbert et al. (2006) showed that across diverse cognitive domains differences in reaction times between experimental and control conditions are differentially associated with the lateral versus medial rostral prefrontal cortex. That is, when performing a meta-analysis the researcher should carefully select the respective experiments, focusing not only on the paradigm name but also check if the process involved in the respective contrast really reflects the critical cognitive process.

In addition to specifying the paradigms for the analysis, inclusion and exclusion criteria need to be specified. There are general criteria that should be applied. These general criteria refer to only including whole brain experiments (see details below) and only including experiments from which coordinates or statistical images in standard anatomical space can be obtained (see details below). For ES-SDM, another general criteria is to only include experiments that report activations and deactivations (or increases and decreases when comparing groups).

Additionally, specific criteria that depend on the particular research question must be specified. Beyond included tasks and paradigms, these specific criteria can relate to analyses and methods. For example, the question might arise if one should only include functional imaging (fMRI) studies (e.g., Kurkela and Dennis, 2016) or studies using either fMRI or positron emission tomography (PET) (e.g., Langner and Eickhoff, 2013; zu Eulenburg et al., 2012).

Examples of other specific inclusion and exclusion criteria relate to aspects of the analysis (e.g. inclusion of only main effects or also of interactions, restricting the meta-analysis to only experiments reporting results on a certain statistical threshold) or to characteristics of the subject group (for example including only healthy subjects or only group comparisons, inclusion of only a specific age range of subjects). Importantly, it should always be kept in mind that the criteria one applies have an impact on how heterogeneous (or homogeneous) the sample of experiments is. Moreover, inclusion and exclusion criteria influence whether or not the sample of experiments is representative for the entire neuroimaging literature available for a specific topic and thus the quality of inclusion. In general, quality of inclusion is given when doing a systematic literature search. However, under certain circumstances it might be limited. For example, when the process investigated and the corresponding inclusion criteria and terminology are defined based on the work of one specific author doing a lot of experiments in