NIDM-Results: a Neuroimaging Data Model to share brain mapping statistical results

Camille Maumet[1], Tibor Auer[2], Alexander Bowring[1], Gang Chen[15], Samir Das[3], Guillaume Flandin[4], Satrajit Ghosh[5], Tristan Glatard[3,6], Krzysztof J. Gorgolewski[7], Karl G. Helmer[8], Mark Jenkinson[9], David Keator[10], B. Nolan Nichols[11], Jean-Baptiste Poline[12], Richard Reynolds[16], Vanessa Sochat[7], Jessica Turner[13] and Thomas E. Nichols[1,14]

[1] Warwick Manufacturing Group, University of Warwick, Coventry, United Kingdom.
[2] MRC Cognition and Brain Sciences Unit, Cambridge, United Kingdom.
[3] McGill Centre for Integrative Neuroscience, Ludmer Centre, Montreal Neurological Institute, Montreal, Quebec, Canada.
[4] Wellcome Trust Centre for Neuroimaging, UCL Institute of Neurology, London, United Kingdom.
[5] McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, USA.
[6] Université de Lyon, CREATIS; CNRS UMR5220; Inserm U1044; INSA-Lyon; Université Claude Bernard Lyon 1, France.
[7] Department of Psychology, Stanford University, Stanford, CA, USA.
[8] Martinos Center for Biomedical Imaging, Massachusetts General Hospital; Dept. of Radiology, Boston, MA, USA.
[9] University of Oxford, UK.
[10] Dept. of Psychiatry and Human Behavior, Dept. of Computer Science, Dept. of Neurology, University of California, Irvine, CA, USA.
[11] Center for Health Sciences, SRI International, Menlo Park, CA, USA.
[12] Helen Wills Neuroscience Institute, H. Wheeler Jr. Brain Imaging Center, University of California at Berkeley, CA, USA.
[13] Psychology and Neuroscience, Georgia State University, Atlanta, GA, USA
[14] Dept. of Statistics, University of Warwick, Coventry, United Kingdom.
[15] Scientific and Statistical Computing Core, National Institute of Mental Health, National Institutes of Health, USA

Introduction

Results of a neuroimaging study are usually shared through the publication of a scientific paper describing the experiment and analysis outcome. While hundreds of gigabytes of data can be generated as part of an functional Magnetic Resonance Imaging (fMRI) experiment, in the literature the authors typically report their results as 1) a list of significant local maxima, i.e. locations in the brain defined in a standard space (e.g. MNI) that pass rigorous statistical testing and 2) a graphical representation of the activations as an image figure.

This practice is unsatisfactory in terms of data re-use as it does not allow for the automatic extraction of acquisition or processing information and it provides only sparse information about the location of the brain activity. And despite the availability of guidelines (Inglis, 2015; Poldrack et al., 2008), ambiguous or incomplete methodological reporting in papers is still commonplace (Carp, 2013). Databases have been built to provide metadata associated with published papers, either manually curated (e.g. BrainMap[1] (Laird et al., 2005)) or automatically-extracted with text-mining algorithms (e.g. NeuroSynth[2] (Yarkoni et al., 2011)), but, ideally, these metadata should be made available by the authors themselves at the time of the publication.

Full representation of neuroimaging results in a machine-readable form would provide unambiguous description and hence support more reproducible and robust science (Button et al., 2013; Carp, 2013). Furthermore, meta-analyses (Kober and Wager, 2010) by providing quantitative syntheses of the rich neuroimaging literature available (>29,000 papers, according to a PubMed search on "fMRI" in Title/Abstract) could help overcome the small sample sizes typically observed in neuroimaging (Carp, 2012). Coordinates typically reported in neuroimaging publications can be synthesized into a meta-analysis using techniques from coordinate-based meta-analysis (e.g. with Activation Likelihood Estimation (Eickhoff et al., 2012), Multilevel Kernel Density Analysis (Kober et al., 2008) or Signed Differential Mapping (Costafreda et al., 2009)). A drawback to these approaches is that peak coordinates lack the level of detail required for optimal meta-analysis. The best practice method is an intensity-based meta-analysis that combines the voxel-wise effect estimates and their standard errors from each study (Salimi-khorshidi et al., 2009). To perform a gold standard meta-analysis, one would therefore need whole-brain 3D image files (statistics, effects sizes maps) from the original analysis, along with metadata about the analysis.

[1] https://brainmap.org/
[2] http://neurosynth.org/

Additional metadata could also describe the experiment itself with further benefit of being able to filter studies by tested hypotheses. Such a level of detail is essential to identify potential confounding factors that are currently being overlooked (e.g. how different smoothing kernels impact the meta-analysis, or the influence of different processing strategies on the outcome of the analysis).

Encouragingly, in the last ten years, the number of publicly available neuroimaging datasets has greatly increased, principally thanks to consortiums dedicated to data sharing such as the Human Connectome Project[3] (Van Essen et al., 2013), the International Data sharing Initiative[4] (Mennes et al., 2013), the Consortium for Reliability and Reproducibility[5] (Zuo et al., 2014), the SchizConnect project[6] (Wang et al., 2016), the Function BioInformatics Research Network project[7] (Zuo et al., 2014), or Study Forrest[8] (Hanke et al., 2014) and to dedicated databases such as OpenfMRI[9] (Poldrack et al., 2013) or NeuroVault[10] (Gorgolewski et al., 2015). Among emerging infrastructures to support data sharing, NeuroVault targets those statistical maps that are necessary for image-based meta-analysis. But, while there is an increasing interest in sharing raw and intermediate derived data (e.g. after motion correction, registration to a template), statistical images supporting neuroimaging publications are still rarely made available.

Data sharing in the neuroimaging community is restrained by a number of psychological and ethical factors that are beyond the scope of the current paper (see (Poldrack and Gorgolewski, 2014; Poline et al., 2012) for a review). But in order to make data sharing common practice, the technological issues that make it challenging must first be addressed.

A key technical issue is the lack of communication between the neuroimaging analysis software packages (e.g. SPM[11] (Penny et al., 2011), FSL[12] (Jenkinson et al., 2012), AFNI[13] (Cox, 1996)) and data sharing tools (e.g. BrainMap, Neurosynth, Neurovault, Brainspell[14]...). While the former are commonly used by most neuroimaging researchers, the latter is still only routinely used by a few specialists. Each of the data sharing tools specified previously use different specifications and formats to store data and metadata about the analysis. Seemingless communication between analysis and data sharing tools is a necessity to lower the barrier to data sharing. A number of metadata pieces are indeed available in a machine-readable form at the time of the analysis but not easily transferrable to the data sharing tools. For example, metadata entered by the user to perform the analysis (e.g. the stimulus onsets for task-based fMRI analysis) and default parameterization used by the analysis software (e.g. the hemodynamic response function) would be highly valuable to be carried forward with the data outputs. If this information is available at the time of the analysis, it is not logical that researchers that are willing to share statistical maps need to manually (re-)enter it. Such a practice is cumbersome, error-prone, and introduces unnecessary barriers to sharing data. To address these barriers entering metadata is often optional (Gorgolewski et al., 2015) but, this strategy makes the missingness issue even more pronounced. To circumvent this issue, extensions have been proposed to automatically extract the metadata from the original analysis files. For example, in the Neurovault database, upload of a full results directory can be performed to automatically load all maps and available metadata. This solution is a step in the right direction to make it easier for the end-user to share data, but can be problematic if the results folder includes data that should not be shared due to privacy issues. Moreover, only limited metadata are currently extracted (e.g. the statistic type: T or F).

A second technical issue is definition of the metadata that should be shared. The space of possible metadata to report is extremely large (encompassing parametrization of acquisition, pre-processing, or

[3] http://www.humanconnectomeproject.org/
[4] http://fcon_1000.projects.nitrc.org/
[5] http://fcon_1000.projects.nitrc.org/indi/CoRR/html/
[6] http://schizconnect.org
[7] http://fbirn-bdr.bic.uci.edu/
[8] http://studyforrest.org
[9] https://openfmri.org/
[10] http://neurovault.org/
[11] http://www.fil.ion.ucl.ac.uk/spm/
[12] http://fsl.fmrib.ox.ac.uk/fsl
[13] http://afni.nimh.nih.gov/
[14] http://brainspell.org/

statistical analysis...) and defining a minimum set is challenging. The optimal set of metadata is highly dependant on the application of interest, and possible applications for shared data are by definition very broad. Further, even for a given application (e.g. meta-analysis) the precise impact of each potential confounding factor is not yet fully understood. While data sharing platforms are the most obvious first contenders to select metadata of interest as part of their implementation, it is beyond the scope of one particular database or software to define what pieces of information should be required for sharing. This leaves an undesired degree of freedom: the choice of metadata and format to encode them.

Finally, there is a level of ambiguity in the way that models and results are described within each neuroimaging software. Mainly due to historical reasons, different neuroimaging software can refer to the same concept using different terms, for example FSL's parameter estimate maps (denoted by pe1.nii.gz, pe2.nii.gz…) are the equivalent of SPM's beta maps (denoted by beta_0001.nii, beta_0002.nii…). The same word might be used to describe different approaches. For instance both SPM and FSL estimates resolution elements (or "resels") but using different approaches. Is it meaningful to report the number of resels in the search space and compare this quantity across software? In order to fully describe an analysis, the sharing of software-specific batch scripts (e.g. SPM matlabbatch files, FSL fsf files, or history stored in AFNI brick headers) would be a simple solution to provide all the parameters from an analysis, but the ability to compare and query across software would still be lacking. Ideally, one should be able to retrieve all studies corresponding to a set of criteria of interest regardless of the software used. This will only be possible if a level of communality can be represented across software. While standard data structure outputs (e.g. a T-Statistic brain map) can be combined despite different software that generated them, the subtle differences in parameters used to generate these standard structures usually go unnoticed. Lack of an ability to standardize and compare these strategies makes it impossible to determine if these subtle differences have broader implications for the results themselves. For example, SPM and FSL are both scaling data to get percent BOLD signal change, but due to differences in how the analysis mask is computed, and due to the different algorithms used to define the mean signal, the data are actually being scaled differently[15]. An added source of variability is the computational environment itself (e.g. (Glatard et al., 2015)). A trade-off needs to be found in the level of details provided, insuring both a level of communality to be able to pull results data across software and outlining enough differences to enable future research.

The evidence above presents a clear need for an automated solution that will enable sharing neuroimaging results using a common descriptive standard across neuroimaging software. NIDM-Results is a data model that can directly address these issues, providing a lightweight solution to share maps generated by neuroimaging studies along with their metadata. While NIDM-Results focuses on mass-univariate studies and is mostly targeted at fMRI, the standard is also suitable for anatomical MRI (with Voxel-Based Morphometry) and Positron Emission Tomography (PET). It was developed under the auspices of the INCF Neuroimaging data sharing Task Force by a core group of experts representing more than 10 labs involved in various facets of neuroimaging (informatics, software development, statistical analysis, development of ontologies…) and in close collaboration with the main neuroimaging software developers. The format is natively implemented in SPM and a Python library is ready for integration in FSL. NeuroVault and CBRAIN are...

Existing work

Encoding of provenance, i.e. keeping track of the processes that were applied to the data encompassing a description of the tools, data flow and workflow parameterization, is a topic of growing interest in science in general. A number of solutions have been proposed in order to support better documentation of research studies. We discuss below the initiatives overlapping our domain of interest: the representation of mass univariate statistical analyses in neuroimaging.

The closest effort to the one presented here is the XML-based Clinical and Experimental Data Exchange (XCEDE) schema (Gadde et al., 2012). XCEDE was developed in the context of the Biomedical Informatics

[15] http://blogs.warwick.ac.uk/nichols/entry/spm_plot_units/

Research Network[16] (BIRN) (Keator et al., 2009) to model information about both the experimental design and results (peaks and clusters) in neuroimaging studies. The XML schema is openly available[17] and was defined to be independent from any particular neuroimaging analysis software. XCEDE has been used by multiple sites across the United States and the United Kingdom in the context of the fBIRN project and is still in use by the Human Imaging Database[18] (Keator et al., 2009). An implementation was provided for SPM[19] (Keator et al., 2006) as well as a set of tools[20]. As part of the fBIRN, XCEDE was used to upload data and metadata to a central database that researchers could then query to retreive data of interest.

As a follow-up to XCEDE, (Keator et al., 2013) introduced the principles of the NeuroImaging Data Model (NIDM), a representation using semantic web technologies and the Resource Description Framework (RDF). Since then, NIDM has been an ongoing effort aiming at thoroughly describing neuroimaging provenance from acquisition to the production of all kinds of derived data[21]. One of the reasons that pushed this group towards RDF was the level of expressivity enabled by graph-based structures and the possibility to form complicated queries across datasets (Keator et al., 2013). The extensibility of semantic web data models, the possibility to interconnect across knowledge domains and to re-use existing constructs are also some of the main strength of this approach. Each class or attribute used in the model is uniquely identified by a URL and associated with a definition (e.g. the SPM neuroimaging software is identified by http://scicrunch.org/resolver/SCR_007037). Terms that are part of a common vocabulary share the same base URL known as a namespace (e.g. http://scicrunch.org/resolver/ is the namespace used for Scicrunch[22]).

A lot of work has been done in the neuroimaging community to provide controlled vocabularies and ontologies defining neuroimaging concepts. The aim of those vocabularies is to provide a list of well-defined terms that can then be referenced as part of a data model to describe a particular aspect of a neuroscience experiment. NeuroLex[23] provides a common platform and already gathers terms from different sources (including previous vocabularies developed by NIF, BIRN…). Interestingly, Neurolex was part of the recent Resource Identification Initiative[24] (RII) (Bandrowski et al., 2016) that publicized the use of those identifiers (e.g. “SCR_007037” for SPM[25]) in research papers. RII is currently focused on the identification of biological resources and has been quickly adopted, with more than 100 journals participating to date (“Anita Bandrowski on Twitter: 708100029609738240,” n.d.).

Beyond controlled vocabularies, ontologies provide a set of terms (and definitions) and describe relations between them (e.g. a ‘Statistic Map’ ‘*was Generated by’* a ‘Contrast Estimation’ activity). Within ontology definitions, two tendencies have emerged. On the one hand, we have ontologies defined as part of a full ecosystem, e.g. OBI[26] and CogPO[27] built upon the OBO foundry, or OntoNeurolog[28] based on the Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) (Masolo et al., 2003). The main advantage of those formal ontologies is the direct interconnection across domains, however this comes at the cost of a steep learning curve. Another tendency in the development of these standards is the definition of smaller scale ontologies with a well-defined limited scope. Examples in neuroimaging include the Cognitive Atlas[29] (Poldrack et al., 2011) for cognitive constructs, which is used, for example, by NeuroVault.

[16] http://www.birncommunity.org/
[17] http://xcede.org/
[18] http://www.birncommunity.org/tools-catalog/human-imaging-database-hid/
[19] http://www.birncommunity.org/tools-catalog/xcede-spm-toolbox/
[20] www.nitrc.org/projects/bxh_xcede_tools/
[21] http://nidm.nidash.org
[22] https://scicrunch.org/
[23] http://neurolex.org/
[24] https://scicrunch.org/resources
[25] https://scicrunch.org/resolver/nif-0000-00343
[26] http://obi-ontology.org/
[27] http://www.cogpo.org/
[28] http://bioportal.bioontology.org/ontologies/ONL-MR-DA, http://bioportal.bioontology.org/ontologies/ONL-DP
[29] http://www.cognitiveatlas.org/

Beyond the neuroimaging domain, representation of mass-univariate neuroimaging results overlaps with more general topics including provenance, statistics, file descriptions… NIDM was built as an extension of the PROV data model[30] (Moreau and Missier, 2013). PROV is a W3C specification that describe provenance in a very generic fashion that is not tied to any domain in particular (cf. (Huynh et al., 2013) for example). PROV defines three types of objects: an *Activity* represents a process that was performed on some data (e.g. the segmentation of an image) and occurred over a fixed period of time; an *Agent* represents someone (human, organization, machine…) that takes responsibility for an activity (e.g. SPM software) and, finally, an *Entity* represents any sort of data, parameters etc. that can be input or output of an activity (e.g. a NIfTI image). PROV also defines a set of relations between those objects (e.g. a segmentation *Activity used* a NIfTI image *Entity*; a segmentation *Activity was associated with* the SPM *Agent* and another NIfTI image *Entity was generated by* the segmentation *Activity*). While other efforts that have been proposed to model provenance such as foundational ontologies such as the OBO foundry or DOLCE. PROV has the advantage to be a lightweight ontology focusing only on provenance and easily extensible to provide domain-specific knowledge.

Standard GLM-type analyses of fMRI data also rely on well-known statistical constructs (e.g. one-sample T-test, two-sample T-test, F-tests, ANOVA, inference, ordinary least squares estimation…) that can be represented beyond the scope of neuroimaging results. The general-purpose STATistics Ontology (STATO) is built on the top of the OBO foundry that aims to provide a set of terms describing statistics.

Material and methods

Modelling

NIDM-Results focuses on mass-univariate models based on a General Linear Model (at the subject or group level). To ease adoption, we restricted the scope of NIDM-Results to metadata that could be automatically extracted and limited user input to metadata that were crucial for the meta-analysis application. This had important practical consequences. Given that pre-processing and statistical analysis are sometimes done using separate pipelines, we focused on the statistical analysis only. The concepts to be represented in NIDM-Results were selected based on (1) meta-analysis best practices; (2) published guidelines to report fMRI studies (Poldrack et al., 2008), and (3), in an effort to ensure continuity with current practice, we also considered the elements displayed as part of results reporting in different neuroimaging software (e.g. peaks, clusters), all those were mitigated by our panel of experts.

When an item, essential for image-based meta-analysis, was not produced as part of the analysis (e.g. the contrast standard error map in SPM) we generated it from existing data. From the guidelines (Poldrack et al., 2008), we looked at the "Statistical modelling" checklist and identified a subset of 12 elements that could be automatically retrieved (cf. Table [table:checklist], first column).

All terms and definitions were thoroughly discussed between our panel of experts in the NIDM working group, which is part of the INCF Neuroimaging Task Force (NIDASH). Since August 2013, we participated in weekly conference calls and 6 focused workshops with a core group of experts representing more than 10 laboratories involved in various facets of neuroimaging (including informatics, software development, statistical analysis, development of ontologies). Furthermore, a separate meeting was organised with each of the development teams of the three major neuroimaging software: SPM, FSL and AFNI to discuss the model and its implementation. Minutes of the meetings and online discussions are publicly available in our shared Google drive[31] and on GitHub under the incf-nidash organization[32].

We defined 3 namespaces: http://purl.org/nidash/nidm# ("nidm:"),http://purl.org/nidash/spm# ("spm:"), http://purl.org/nidash/fsl# ("fsl:"). Anything that could be represented across software or was a generic concept was defined in "nidm:" namespace. Software-specific namespaces: ("spm:", "fsl:") were reserved for

[30] https://www.w3.org/TR/prov-dm/
[31] https://drive.google.com/open?id=0B-BLof5_SOh8bWR3UDE4WTdELXM
[32] https://github.com/incf-nidash/nidm

the description of software idiosyncrasies (e.g. global null inference for conjunction testing in SPM). It is important to note that terms in the "nidm:" namespace can correspond to different approaches in different neuroimaging software (and varying implementation).

For each piece of information, we checked if an appropriate term was available in publicly available ontologies: in particular STATO for statistics term, PROV for provenance but also to a lesser extent Dublin Core and Neurolex/RRID. We engaged with STATO by means of GitHub issues[33] to propose new terms when needed. When no term was found to describe a given neuroimaging concept of interest, we created a new term and carefully crafted a definition.

Implementation

In an effort to make NIDM-Results available to a majority of neuroimaging researchers, we created a set of exporter for SPM and FSL. We also engaged with NeuroVault to propose a one-click upload of NIDM-Results archives.

Experiments

A simple image-based meta-analysis

From 25 pain studies exported using NIDM-Results we performed an image-based meta-analysis with the gold standard approach (MFX 3rd level) using FSL's FLAME 1 (Smith et al., 2001). We corrected for difference in data scaling between software by rescaling the FSL maps to a target intensity of 100 (instead of 10 000 by default).

Reporting of neuroimaging results

From a study exported with NIDM-Results we wrote a script to extract the information of interest to describe group statistics. The paragraph that was generated could, for instance, be used as part of the method section in a research paper.

Examples and queries

To the attention of developers willing to build application using NIDM-Results, we exported 10 fMRI experiments (using OpenfMRI datasets) and made them publicly available. As part of our specification, we also provided a set of queries that can be used to retrieve information of interest.

Results

Overview of the model

An overview of NIDM-Results is proposed in fig. 2. A bundle, (i.e. a "super-entity") provides a version number to facilitate further evolutions of the model. An agent represents the exporter that generated the bundle. Within the bundle, the remaining of the model is made of:

- 3 activities corresponding to the typical steps of statistical hypothesis testing: *model parameters estimation*, *contrast estimation* and *inference*.
- 28 entities and;
- 1 agent (the neuroimaging software).

The statistical model is defined by the *design matrix* and *error model* entities that are both used by *model parameters estimation*. *Data scaling* describes the scaling applied to the data before model fitting, this is especially relevant for first-level fMRI experiments. A set of *parameter estimate maps* is generated by

[33] https://github.com/ISA-tools/stato/issues?utf8=%E2%9C%93&q=is%3Aissue+nidm+

model parameters estimation along with the analysis *mask*, a *residual mean squares map* and a *grand mean map* that can be used to check the performance of the data scaling.

The *contrast estimation* activity uses a subset of the *parameter estimates*, the *residual mean squares map* and the analysis *mask* to generate a *statistic map*. For T-tests, a *contrast map* along with its *standard error map* are also created while for F-tests a *contrast explained mean square map* (i.e. the numerator of an F-statistic) is provided.

For each map entity (e.g. *contrast map*, *statistic map*…), except the *parameter estimate maps*, the corresponding nifti image is included in the export and linked to the entity using a *prov:atLocation* attribute (cf. 1 for an example).

```
@prefix prov: <http://www.w3.org/ns/prov#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix nfo: <http://www.semanticdesktop.org/ontologies/2007/03/22/nfo#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix crypto: <http://id.loc.gov/vocabulary/preservation/cryptographicHashFunctions#> .

@prefix nidm_StatisticMap: <http://purl.org/nidash/nidm#NIDM_0000076> .
@prefix nidm_statisticType: <http://purl.org/nidash/nidm#NIDM_0000123> .
@prefix nidm_contrastName: <http://purl.org/nidash/nidm#NIDM_0000085> .
@prefix nidm_effectDegreesOfFreedom: <http://purl.org/nidash/nidm#NIDM_0000091> .
@prefix nidm_inCoordinateSpace: <http://purl.org/nidash/nidm#NIDM_0000104> .
@prefix nidm_errorDegreesOfFreedom: <http://purl.org/nidash/nidm#NIDM_0000093> .
@prefix obo_tstatistic: <http://purl.obolibrary.org/obo/STATO_0000176> .

niiri:statistic_map_id a prov:Entity , nidm_StatisticMap: ;
   rdfs:label "Statistic Map: listening > rest" ;
   prov:atLocation "TStatistic.nii.gz"^^xsd:anyURI ;
   nidm_statisticType: obo_tstatistic: ;
   nfo:fileName "TStatistic.nii.gz"^^xsd:string ;
   dct:format "image/nifti"^^xsd:string ;
   nidm_contrastName: "listening > rest"^^xsd:string ;
   nidm_effectDegreesOfFreedom: "1"^^xsd:float ;
   nidm_inCoordinateSpace: niiri:coordinate_space_id_1 ;
   crypto:sha512 "e43b6e01b0463fe7d40782137867ae43b6e01b0463fe7d40782137867a"^^xsd:string ;
   prov:wasGeneratedBy niiri:contrast_estimation_id;
   nidm_errorDegreesOfFreedom: "72.9999999990787"^^xsd:float .
```

Fig. 1. Representation of a Statistic Map entity in turtle. The prov:atLocation attribute specifies the path to the corresponding NIfTI image.

Finally, the *inference* activity uses a *statistic map* and generates an *excursion set map* given a *height* and an *extent thresholds*. The *peak definition criteria* and *cluster definition criteria* entities, used by *inference*, provides the connectivity criterion and minimal distance between peaks (e.g. default is set to 8 mm for SPM and 0 mm for FSL). The *inference* activity can be replaced by a *conjunction inference* which uses more than one statistic map. An optional *display mask map* entity can be used to represent contrast masking, i.e. to restrict the display but do not affect the correction for multiple comparisons. The *inference* activity also generates the *search space map* that represents the search region in which the inference was performed (i.e. the intersection of all input mask maps, except for the display mask map).

A set of *supra-threshold clusters* is derived from the *excursion set map* and a set of *peaks* is derived from each *cluster*. Those are the clusters and peaks that are typically reported as a results of a neuroimaging study.

A software agent represent the analysis software used to compute the analysis. This agent is responsible for all activities within the bundle.

Each activity, entity and agent has a number of predefined attributes. For instance, the list of attributes of a *contrast map* is provided in fig. 3. In total, 340 terms were defined (147 classes, 154 attributes, 39 individuals). For more details the interested reader is referred to the specification available at: http://nidm.nidash.org/specs/nidm-results.html. The model is also defined as an owl file available at: https://github.com/incf-nidash/nidm.

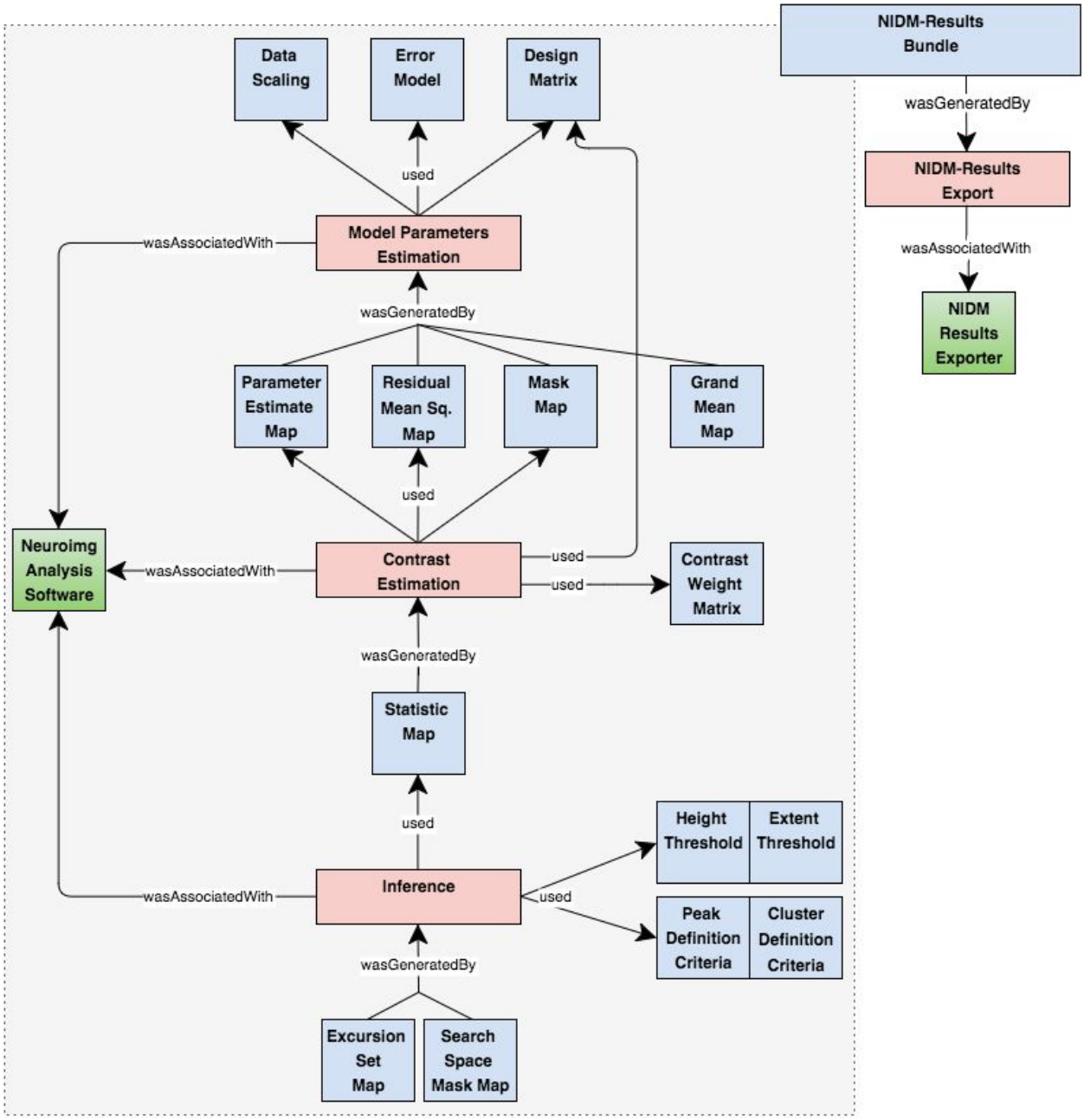

Fig. 2. NIDM-Results core structures. Each box represents an Activity (red), Entity (blue) or Agent (green) as defined by PROV.

nidm:'Contrast Map' ☆+: A map whose value at each element (e.g., pixel, voxel, vertex, or face) is a contrast estimate.

nidm:'Contrast Map' is a nidm:'Map' used by nidm:'Inference' and generated by nidm:'Contrast Estimation'.

A nidm:'Contrast Map' has attributes:

- rdfs:label: (OPTIONAL) Human readable description of the nidm:'Contrast Map'.
- crypto:sha512: (OPTIONAL) Secure Hash Algorithm 512.
- dct:format: (OPTIONAL) The file format, physical medium, or dimensions of the resource. (range xsd:string).
- nidm:'contrast Name': (OPTIONAL) Name of the contrast. (range xsd:string).
- nidm:'has Map Header': (OPTIONAL) A Property that associates an additional file containing the 'Map Header' with a map. (range nidm:'Map Header').
- nidm:'in Coordinate Space': (OPTIONAL) Property of a DataArray to associate a 'Coordinate Space' with a physical file. (range nidm:'Coordinate Space').
- nfo:'fileName': (OPTIONAL) Name of the file, together with the extension. (range xsd:string).
- prov:atLocation: (OPTIONAL) The Location of any resource.

Fig 3. Description of a Contrast Map entity (excerpt of the NIDM-Results 1.2.0 specification[34]).

Demonstration, how-tos

NIDM-Results export from SPM results

SPM12 natively supports export to NIDM-Results through a contextual menu in the results table or non-interactively via the batch interface as illustrated in fig 4.

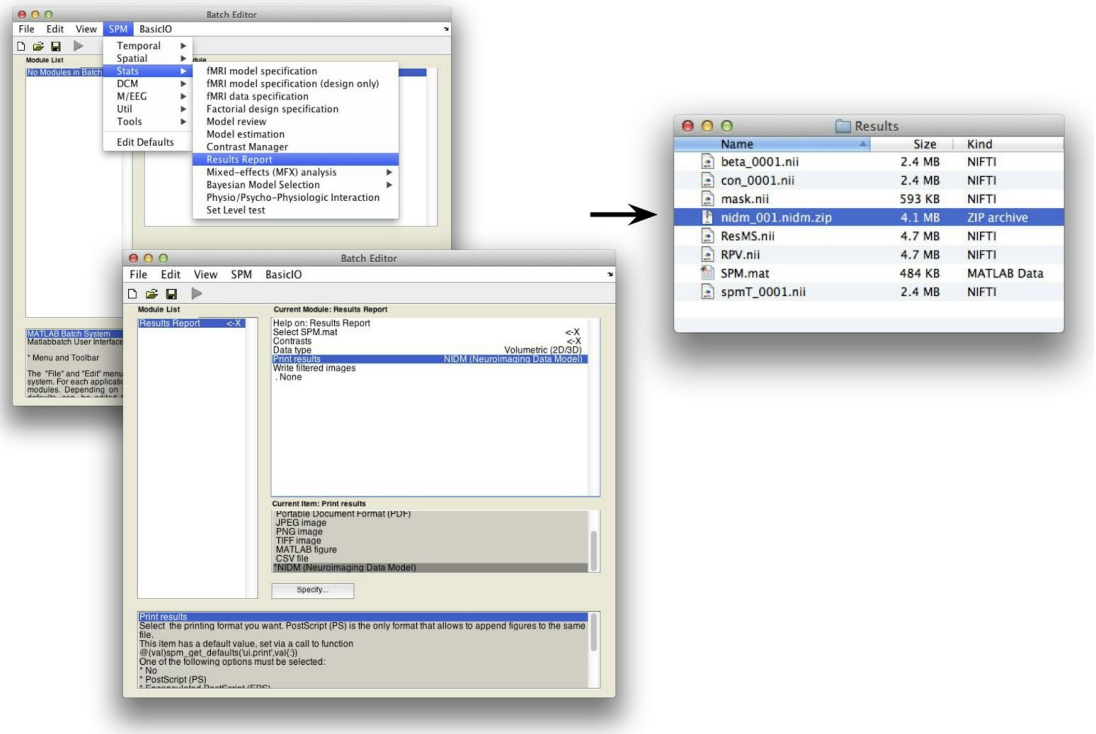

Fig. 4: NIDM-Results export in SPM12

[34] http://nidm.nidash.org/specs/nidm-results_120.html

NIDM-Results export from FSL results

Export of feat analyses from FSL to NIDM-results can be performed using the python module `nidmfsl` as illustrated in fig. 5. `nidmfsl` is available on the Python Package Index[35] and can therefore be installed using the following command:

```
pip install nidmfsl
```

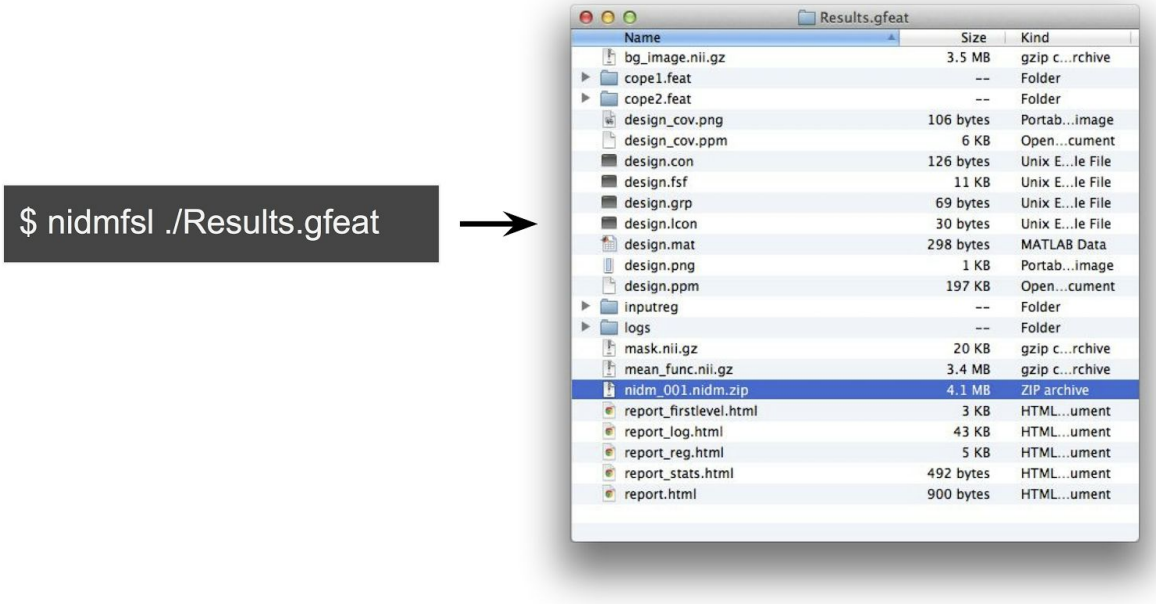

Fig. 5. NIDM-Results export in FSL v5.0.

NIDM-Results export from CBRAIN and NeuroVault

`nidmfsl` was integrated as a plugin[36] of the CBRAIN web platform for high-performance computing (Sherif et al., 2014). As a result, any FSL feat analysis performed in CBRAIN can be exported to NIDM-Results.

NIDM-Results use cases

A simple meta-analysis

The Python script used to run the meta-analysis on the NIDM-Results export is available on GitHub[37]. Fig. 6 provides a schematic overview of the different steps.

First, each NIDM-Results export was queried to retrieve the image data needed for the meta-analysis (i.e. the contrast image and contrast standard error image) along with the analysis mask. The name of the corresponding contrast was associated to each map to allow for the selection of the appropriate contrast. The neuroimaging software used for the analysis was also extracted in order to identify which study estimates would need re-scaling.

Second, the contrast and standard error estimates were selected according to the contrast name, re-scaled if needed and combined in a mixed-effects GLM. Areas of significant activation ($p<0.05$ voxel-wise

[35] https://pypi.python.org/pypi
[36] https://github.com/glatard/cbrain-plugins-nidm
[37] https://github.com/incf-nidash/nidmresults_paper/blob/master/simple_meta_analysis.py

FDR-corrected) found by the pain meta-analysis are displayed in fig. 7. These are in lines with previous results from the literature (e.g. (Salimi-khorshidi et al., 2009)).

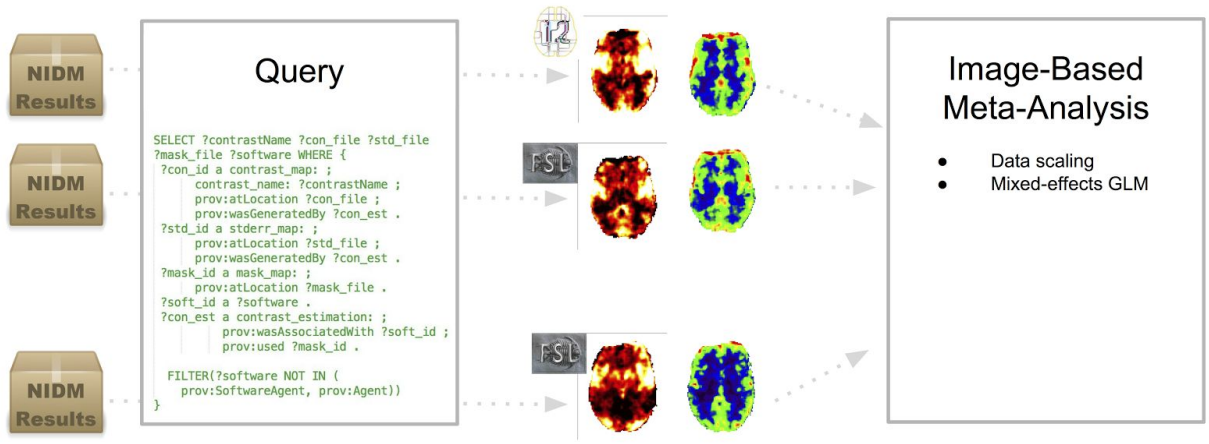

Fig. 6. Pain meta-analysis using NIDM-Results. Each NIDM-Results archive is queried to retrieve the contrast map, the contrast standard error map, the software used and the analysis mask, appropriate studies are selected according to the contrast name. The study maps are combined in a mixed-effects GLM.

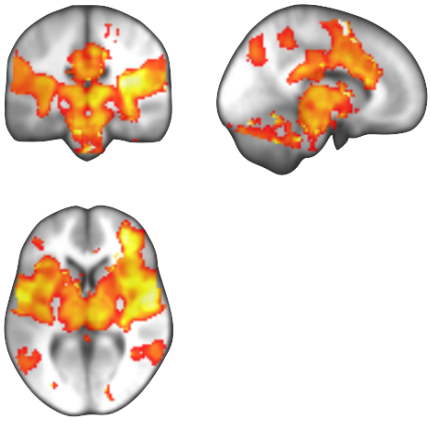

Fig. 7. Pain meta-analysis: areas of significant activation (p<0.05 voxel-wise FDR-corrected).

Reporting of neuroimaging results

The Python script used to create the report based on a NIDM-Results export is available on GitHub[38]. The SPARQL query used to retrieve the information of interest for each report is displayed in fig. 8.

[38] https://github.com/incf-nidash/nidmresults_paper/blob/master/query_results_report.py

Fig. 8. Reporting group statistics: query.

Discussions

As for any new format, gaining acceptance within the neuroimaging community will be crucial for NIDM-Results. To insure a level of consensus, including the point-of-view of different actors in the field, NIDM-Results was built as part of a collaborative effort. More feedback is welcomed from the community and the standard might evolve in the future. All comments are welcome. We also made a strong commitment to make implementations available. Taking advantage of the fact that most GLM-based neuroimaging studies are performed using a limited number of software (> 75% for SPM or FSL, > 90% for SPM, FSL or AFNI according to (Carp, 2012)), we developed implementation for SPM and FSL. We are currently working with AFNI developers to further extend the coverage of NIDM-Results export.

While we focused on implementation at the level of the neuroimaging software, the development of applications processing NIDM-Results is also crucial; in particular to act as incentives for the neuroimagers to use this format. As an example of application, we liaised with NeuroVault to propose a one-click upload of NIDM-Results. Here, users can benefit from all Neurovault features including state-of-the-art visualisations but also sharing, either privately or publicly depending on the stage of the project. This process can ease communication between researchers working on different platforms or used to a different set of neuroimaging tools.

As a first step to provide machine-readable metadata, we restricted our scope to automatically extractable pieces of information. This limited the amount of information that could be represented. For instance, the description of the paradigm was limited to a CSV file and an image of the design matrix plus a list of regressor names. Ideally, to be able to automatically query for studies of interest, one would need a more thorough description of the paradigm and of the cognitive constructs involved. While vocabularies are becoming available (e.g. Cognitive Atlas, CogPO), description of fMRI paradigms is still a topic of active research. Some level of manual interaction is therefore still needed to compute a meta-analyses based on

NIDM-Results exports. Nevertheless, NIDM-Results allows for the automation of part of the meta-analysis as described in our results. In the future, as a consensus develops on the description of paradigms, NIDM-Results could easily be extended to include this information.

The model is mainly focused on parametric analysis of BOLD fMRI. In practice, any type of mass-univariate GLM-based study, e.g. VBM experiments could be encoded. Thank to the intrinsic extensibility of RDF models, variants could be proposed. For example, an extension for non-parametric statistics is under discussion[39].

NIDM-Results is part of a broader effort (NIDM) that aims at representing different levels of provenance in neuroimaging experiments. Other NIDM efforts include: NIDM-Experiment that targets the representation of raw data and information on the participants and NIDM-Workflows which focuses on detailed description of data analysis parametrization. While those efforts are still under development, our goal is to keep a link between those components to eventually provide a complete representation of neuroimaging provenance.

NIDM-Results, as a lightweight machine-readable summary of a GLM-based neuroimaging study, also opens the field to new applications. For instance, in the future, one could imagine extending the classical research paper by providing alongside a machine-readable description of the experiment in form of a NIDM-Results document.

Conclusion

We have introduced NIDM-Results a data model to encode the results of a GLM-based neuroimaging studies. Export as a NIDM-Result archive is available in SPM12 (natively) and in FSL 5.0. Discussions with AFNI development team are in-progress to extend NIDM-Results coverage. NIDM-Results archives can be easily be uploaded to Neurovault providing access to a common platform to share and visualize neuroimaging results. In future work, extensions of the model will be provided to encode voxel-wise non-parametric studies. As a machine-readable representation of neuroimaging results, NIDM-Results is a step forward to support reproducibility of neuroimaging studies.

Acknowledgments

We gratefully acknowledge Matthew Webster, Paul McCarthy, Eugene Duff and Steve Smith, from the FMRIB and Robert Cox and Ziad Saad from the NIH, for their inputs on the integration of NIDM-Results within FSL and AFNI; as well as the NIDASH task force members for their inputs during the development of the model. The INCF supported and organised the task force meetings in which the model was discussed. We also gratefully acknowledge the Tracey group at FMRIB for sharing their pain datasets used in the meta-analysis. CM and TEN were supported by the Wellcome Trust. TA was supported by the Medical Research Council (United Kingdom) [MC-A060-53114].

References

Anita Bandrowski on Twitter: 708100029609738240. (n.d.). *Twitter*. Retrieved March 29, 2016, from https://twitter.com/Anitabandrowski/status/708100029609738240

Bandrowski, A., Brush, M., Grethe, J. S., Haendel, M. A., Kennedy, D. N., Hill, S., Hof, P. R., et al. (2016). The Resource Identification Initiative: A cultural shift in publishing. *The Journal of comparative neurology*, *524*(1), 8–22.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., and Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature reviews. Neuroscience*, *14*(5), 365–376. Nature Publishing Group.

Carp, J. (2012). The secret lives of experiments: methods reporting in the fMRI literature. *NeuroImage*, *63*(1), 289–300. Elsevier Inc.

Carp, J. (2013). Better living through transparency: improving the reproducibility of fMRI results through

[39] https://github.com/incf-nidash/nidm/pull/233

comprehensive methods reporting. *Cognitive, affective & behavioral neuroscience*, *13*(3), 660–666.
Costafreda, S. G., David, A. S., and Brammer, M. J. (2009). A parametric approach to voxel-based meta-analysis. *NeuroImage*, *46*(1), 115–122.
Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and biomedical research, an international journal*, 162–173.
Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., and Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage*, *59*(3), 2349–2361.
Gadde, S., Aucoin, N., Grethe, J. S., Keator, D. B., Marcus, D. S., and Pieper, S. (2012). XCEDE: an extensible schema for biomedical data. *Neuroinformatics*, *10*(1), 19–32.
Glatard, T., Lewis, L. B., Ferreira da Silva, R., Adalat, R., Beck, N., Lepage, C., Rioux, P., et al. (2015). Reproducibility of neuroimaging analyses across operating systems. *Frontiers in neuroinformatics*, *9*, 12.
Gorgolewski, K. J., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S. S., Maumet, C., Sochat, V. V., et al. (2015). NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Frontiers in neuroinformatics*, *9*, 8.
Hanke, M., Baumgartner, F. J., Ibe, P., Kaule, F. R., Pollmann, S., Speck, O., Zinke, W., et al. (2014). A high-resolution 7-Tesla fMRI dataset from complex natural stimulation with an audio movie. *Scientific data*, *1*, 140003.
Huynh, T. D., Groth, P., and Zednik, S. (2013). PROV Implementation Report. (T. D. Huynh, P. Groth, and S. Zednik, Eds.). World Wide Web Consortium. Retrieved from http://eprints.soton.ac.uk/358440/
Inglis, B. (2015). A checklist for fMRI acquisition methods reporting in the literature. *The Winnower*. doi: 10.15200/winn.143191.17127
Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., and Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790.
Keator, D. B., Gadde, S., Grethe, J. S., Taylor, D. V., and Potkin, S. G. (2006). A general XML schema and SPM toolbox for storage of neuro-imaging results and anatomical labels. *Neuroinformatics*, *00*, 199–211.
Keator, D. B., Helmer, K., Steffener, J., Turner, J. A., Van Erp, T. G. M., Gadde, S., Ashish, N., et al. (2013). Towards structured sharing of raw and derived neuroimaging data across existing resources. *NeuroImage*, *82*, 647–661. Elsevier B.V.
Keator, D. B., Wei, D., Gadde, S., Bockholt, J., Grethe, J. S., Marcus, D., Aucoin, N., et al. (2009). Derived Data Storage and Exchange Workflow for Large-Scale Neuroimaging Analyses on the BIRN Grid. *Frontiers in neuroinformatics*, *3*, 30.
Kober, H., Barrett, L. F., Joseph, J., Bliss-Moreau, E., Lindquist, K., and Wager, T. D. (2008). Functional grouping and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *NeuroImage*, *42*(2), 998–1031.
Kober, H., and Wager, T. D. (2010). Meta-analysis of neuroimaging data. *Wiley interdisciplinary reviews. Cognitive science*, *1*(2), 293–300.
Laird, A. R., Lancaster, J. L., and Fox, P. T. (2005). BrainMap: the social evolution of a human brain mapping database. *Neuroinformatics*, *3*(1), 65–78.
Masolo, C., Borgo, S., Gangemini, A., Guarino, N., Oltramari, A., and Schneider, L. (2003). The wonderweb library of fundational ontologies and the dolce ontology. wonderweb deliverable d18, final report (vr. 1.0. 31-12-2003). *The WonderWeb Library of Fundational Ontologies and the DOLCE ontology. WonderWeb Deliverable D18, Final Report (vr. 1. 0. 31-12-2003)*.
Mennes, M., Biswal, B. B., Castellanos, F. X., and Milham, M. P. (2013). Making data sharing work: the FCP/INDI experience. *NeuroImage*, *82*, 683–691. Elsevier Inc.
Moreau, L., and Missier, P. (2013, April). *PROV-DM: The PROV Data Model*. World Wide Web Consortium. Retrieved from http://eprints.soton.ac.uk/356851/
Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., and Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images: the analysis of functional brain images*. Academic press.
Poldrack, R. a., Barch, D. M., Mitchell, J. P., Wager, T. D., Wagner, A. D., Devlin, J. T., Cumba, C., et al. (2013). Toward open sharing of task-based fMRI data: the OpenfMRI project. *Frontiers in neuroinformatics*, *7*(July), 12–12.

Poldrack, R. A., Fletcher, P. C., Henson, R. N., Worsley, K. J., Brett, M., and Nichols, T. E. (2008). Guidelines for reporting an fMRI study. *NeuroImage*, *40*(2), 409–414.

Poldrack, R. A., and Gorgolewski, K. J. (2014). Making big data open: data sharing in neuroimaging. *Nature neuroscience*, *17*(11), 1510–1517.

Poldrack, R. a., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., Parker, D. S., et al. (2011). The cognitive atlas: toward a knowledge foundation for cognitive neuroscience. *Frontiers in neuroinformatics*, *5*(September), 17–17.

Poline, J.-B., Breeze, J. L., Ghosh, S., Gorgolewski, K., Halchenko, Y. O., Hanke, M., Haselgrove, C., et al. (2012). Data sharing in neuroimaging research. *Frontiers in neuroinformatics*, *6*(April), 9–9.

Salimi-khorshidi, G., Smith, S. M., Keltner, J. R., Wager, T. D., and Nichols, T. E. (2009). Meta-analysis of neuroimaging data: a comparison of image-based and coordinate-based pooling of studies. *NeuroImage*, *45*(3), 810–823. Elsevier B.V.

Sherif, T., Rioux, P., Rousseau, M.-E., Kassis, N., Beck, N., Adalat, R., Das, S., et al. (2014). CBRAIN: a web-based, distributed computing platform for collaborative neuroimaging research. *Frontiers in neuroinformatics*, *8*, 54.

Smith, S., Bannister, P. R., Beckmann, C., Brady, M., Clare, S., Flitney, D., Hansen, P., et al. (2001). FSL: New tools for functional and structural brain image analysis. *NeuroImage*, *13*(6), 249. Academic Press.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., and WU-Minn HCP Consortium. (2013). The WU-Minn Human Connectome Project: an overview. *NeuroImage*, *80*, 62–79.

Wang, L., Alpert, K. I., Calhoun, V. D., Cobia, D. J., Keator, D. B., King, M. D., Kogan, A., et al. (2016). SchizConnect: Mediating neuroimaging databases on schizophrenia and related disorders for large-scale integration. *NeuroImage*, *124*(Pt B), 1155–1167.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., Wager, T. D., and Essen, D. C. V. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature methods*, *8*(8), 665–670.

Zuo, X.-N., Anderson, J. S., Bellec, P., Birn, R. M., Biswal, B. B., Blautzik, J., Breitner, J. C. S., et al. (2014). An open science resource for establishing reliability and reproducibility in functional connectomics. *Scientific data*, *1*, 140049.

Zuo, Xi-Nian, Jeffrey S. Anderson, Pierre Bellec, Rasmus M. Birn, Bharat B. Biswal, Janusch Blautzik, John C. S. Breitner, et al. 2014. “An Open Science Resource for Establishing Reliability and Reproducibility in Functional Connectomics.” *Scientific Data* 1 (December): 140049.