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, 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

Abstract

Only a tiny fraction of the data and metadata produced by an fMRI study is finally conveyed to the community. This lack of transparency not only hinders the reproducibility of neuroimaging results but also impairs future meta-analyses. In this work we introduce NIDM-Results, a standard providing a machine-readable description of neuroimaging statistical results along with key image data summarising the experiment. NIDM-Results provides a unified representation of mass univariate analyses including a level of detail consistent with available best practices. This standard allows authors to relay methods and results in a standard format that is not tied to a particular neuroimaging software package. Tools are available to export NIDM-Result graphs and associated files from the widely used SPM and FSL software packages, and the NeuroVault repository can import NIDM-Results archives. The specification is publically available at: http://nidm.nidash.org/specs/nidm-results.html.

Introduction

A neuroimaging technique like functional Magnetic Resonance Imaging (fMRI) generates hundreds of gigabytes of data, yet only a tiny fraction of that information is finally conveyed to the community. In a typical paper, the entire results report consists of 1) a list of significant local maxima, i.e. locations in the brain defined in a standard atlas space inferred to be distinct from noise, and 2) a graphical representation of the activations as an image figure.

This practice is unsatisfactory for three reasons. First, because it represents a massive loss of information from the raw and even the derived data used to draw the conclusion of the study. For example, a meta-analysis (in settings other than neuroimaging) combines estimates of effects of interest and their uncertainty across studies. In brain imaging, the locations of local maxima have no measures of uncertainty reported. While neuroimaging meta-analysis methods for coordinate data exist (Eickhoff et al. 2012; Kober et al. 2008; Costafreda, David, and Brammer 2009) they are a poor approximation to the meta-analysis that would be obtained if the image data were available (Salimi-khorshidi et al. 2009). Even though there are emerging infrastructures to support sharing of neuroimaging data (e.g. NeuroVault[1] (K. J. Gorgolewski et al. 2015)), these are still rarely utilised.

[1] http://neurovault.org/

Second, despite the availability of guidelines (R. A. Poldrack et al. 2008; Inglis 2015; Nichols et al. 2016), ambiguous or incomplete methodological reporting in papers is still commonplace (Carp 2013) hindering the robustness and reproducibility of scientific results (Button et al. 2013; Carp 2013).

Finally, key methodological details of the study are described in free-form text in a paper and not available in machine-readable form, making these metadata essentially unsearchable. Databases have been built to provide metadata associated with published papers, either manually curated (e.g. BrainMap[2] (Laird, Lancaster, and Fox 2005)) or automatically-populated using text-mining algorithms (e.g. NeuroSynth[3] (Yarkoni et al. 2011)), but, ideally, these metadata should be made available by the authors themselves at the time of the publication, together with the data. Additionally, searchable metadata, could help 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).

In order to make neuroimaging results available in a machine-readable form a number of key technical issues have to be addressed. First, the scope of the metadata to be shared must be defined. The space of possible metadata to report is extremely large encompassing experimental design, acquisition, pre-processing, statistical analysis, etc. The optimal set of metadata is highly dependant on the application of interest and possible applications of shared data are broad. For example, in a meta-analysis, the contrast standard error map is required, while a comparison across neuroimaging processing pipelines would require a complete description of the analysis pipeline including software-specific parameterization.

Another technical issue is the need to define a common representation across neuroimaging software packages. While the three main neuroimaging software packages, SPM[4] (Penny et al. 2011), FSL[5] (Jenkinson et al. 2012) and AFNI[6] (Cox 1996), all implement similar analyses, they often use different terms to refer to the same concept. For example, FSL's parameter estimate maps (e.g. pe1.nii.gz) are the equivalent of SPM's beta maps (e.g. beta_0001.nii). They can also use the same term when referring to different concepts. For example, SPM and FSL both use a global scaling of the data to get "percent BOLD signal change", but due to differences in how the brain mask and mean signal are computed, the data are scaled quite differently[7] and are not comparable. 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. Pipeline systems like NiPype (K. Gorgolewski et al. 2011), LONI Pipeline (Rex, Ma, and Toga 2003) and CBRAIN (Sherif et al. 2014) do explicitly model analysis steps, but a large volume of research is still conducted directly with tools not embedded in pipelines. Ideally, one should be able to identify all studies corresponding to a set of criteria of interest regardless of the software used. This will only be possible if information about results across software can be represented using common data elements and structures.

A machine-readable representation of neuroimaging data and results, using a common descriptive standard across neuroimaging software packages, would address these issues of comparability and transparency.

A previous effort in this direction was the XML-based Clinical and Experimental Data Exchange (XCEDE) schema (Gadde et al. 2012), developed in the context of the Biomedical Informatics Research Network[8] (BIRN) (David B. Keator et al. 2009). XCEDE modelled information about

[2] https://brainmap.org/
[3] http://neurosynth.org/
[4] http://www.fil.ion.ucl.ac.uk/spm/
[5] http://fsl.fmrib.ox.ac.uk/fsl
[6] http://afni.nimh.nih.gov/
[7] http://blogs.warwick.ac.uk/nichols/entry/spm_plot_units/
[8] http://www.birncommunity.org/

both the experimental design and results (peaks and clusters) in neuroimaging studies. This XML schema was defined to be independent from any particular neuroimaging analysis software and was made openly available[9]. 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[10] (David B. Keator et al. 2009). An implementation was provided for SPM[11] (D. B. Keator et al. 2006) as well as a set of tools[12]. However, the XCEDE model was not implemented by other imaging software, supported limited provenance information, and did not offer the ability to jointly share image data summarising the experiment.

Beyond neuroimaging, 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. Among them, the PROV data model[13] (Moreau and Missier 2013) is a W3C specification to describe provenance on the web. PROV is defined in a generic fashion that is not tied to a domain in particular (cf. (Huynh, Groth, and Zednik 2013) for examples of implementations).

The NeuroImaging Data Model (NIDM) (D. B. Keator et al. 2013)[14] was created to expand upon the initial development of XCEDE, introducing a domain-specific extension of PROV using semantic web technologies and the Resource Description Framework (RDF). The goal of NIDM is to provide a complete description of provenance for neuroimaging studies, from raw data to the final results including all the steps in-between. The core motivation of NIDM is to support data sharing and data reuse in neuroimaging by providing rich machine-readable metadata. Since its first developments in 2011, NIDM has been an ongoing effort and is currently comprised of three complementary projects: NIDM-Experiment, NIDM-Workflows and NIDM-Results. NIDM-Experiment targets the representation of raw data generated by the scanner and information on the participants. NIDM-Workflows focuses on the description of data analysis parameterization, including detailed software-specific variations. NIDM-Results, presented here, deals with the representation of mass-univariate neuroimaging results using a common descriptive standard across neuroimaging software packages.

A motivating use case for NIDM-Results was neuroimaging meta-analysis, but the format also produces a detailed machine-readable report of many facets of an analysis. The implementation of NIDM-Results within SPM and FSL, two of the main neuroimaging software packages, provides an automated 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 International Neuroinformatics Coordinating Facility (INCF) Neuroimaging data sharing Task Force (NIDASH) which comprises a core group of experts representing more than ten labs involved in various facets of neuroimaging (including statistical analysis, informatics, software development, ontologies). It also involved close collaboration with the main neuroimaging software developers. The format is natively implemented in SPM and a NIDM-Results exporter is available for FSL and will be integrated in a future version of FSL. Both NeuroVault and CBRAIN support export to NIDM-Results and NeuroVault additionally can import NIDM-Results archives.

Results

[9] http://xcede.org/
[10] http://www.birncommunity.org/tools-catalog/human-imaging-database-hid/
[11] http://www.birncommunity.org/tools-catalog/xcede-spm-toolbox/
[12] www.nitrc.org/projects/bxh_xcede_tools/
[13] https://www.w3.org/TR/prov-dm/
[14] http://nidm.nidash.org

Model

Definitions

The definitions provided below are used throughout the manuscript:

- **NIDM-Results graph**: A particular instance of a representation of data and metadata complying with the NIDM-Results specification.
- **NIDM-Results serialization**: A text file rendering of a NIDM-Results graph.
- **NIDM-Results pack**: A compressed file containing a NIDM-Results serialization and some or all of the referenced image data files.

Overview

The NIDM-Results standard is defined by a W3-style specification, publicly available at http://nidm.nidash.org/specs/nidm-results.html and by an ontology (owl) file available on Bioportal[15]. It is comprised of a controlled vocabulary, as well as instructions of how to use PROV to represent mass-univariate neuroimaging results. The model provides terms to describe key elements of neuroimaging methods using a common framework across neuroimaging software packages. For example, as illustrated in Fig. 1, error models are described in terms of assumed variance (homoscedastic, heteroscedastic) and assumed covariance structure (independent, spatially correlated, etc.) and how these structures vary in space (defined independently at each voxel, globally throughout the brain or spatially regularised).

The current version, NIDM-Results 1.3.0, defines 214 terms (140 classes and 74 attributes) of which 45 terms are re-used from external vocabularies and ontologies. All terms are defined as specialisations of the PROV terms. Three namespaces are defined: http://purl.org/nidash/nidm# ("nidm:"), http://purl.org/nidash/spm# ("spm:") and http://purl.org/nidash/fsl# ("fsl:"). Anything that could be represented across software or that is a generic concept is defined in the "nidm:" namespace. Software-specific namespaces: ("spm:", "fsl:") are reserved for the description of functionality unique to one software (e.g. global null inference for conjunction testing in SPM).

Fig. 2 provides an overview of NIDM-Results. In the description below, terms in single quote correspond to elements defined by the model, identifiers for those terms are provided in Table 1.

The main entity is a *'NIDM-Results bundle'*, a specialisation of a 'Bundle' as defined in PROV, i.e. an entity gathering a set of entities, activities and agents. A *'NIDM-Results bundle'* contains a description of the mass univariate results provenance and is typically made up of:

- 3 activities representing the main steps of statistical hypothesis testing: '*model parameter estimation'*, '*contrast estimation'* and '*inference'*.
- 26 types of entities (of which 6 are optional) representing inputs and outputs of the activities;
- 3 agents representing the '*neuroimaging analysis software'*, the *'person'* or *'study group population'* who participated in the study and the type of *'imaging instrument'* used.

The statistical model is described in the '*design matrix'* and '*error model'* entities that are both used by the '*model parameter estimation'* activity. The *'data'* entity describes the scaling applied to the data before model fitting (especially relevant for first-level fMRI experiments) and links to the participants (as a 'person' or a group) and the 'imaging instrument' used to acquire the data (e.g. a magnetic resonance imaging scanner or an electroencephalography machine). A set of '*parameter estimate map's* is generated by the '*model parameter estimation'* activity along with the analysis '*mask map'*, a '*residual mean squares map'* and a '*grand mean map'* that can be used to check the performance of the data scaling. Optionally, a *'resels per voxel map'* can also be generated to record local variations in noise smoothness.

[15] http://bioportal.bioontology.org/ontologies/NIDM-RESULTS

The '*contrast estimation'* activity uses a subset of the '*parameter estimate maps'*, the '*residual mean squares map'* and the analysis '*mask map'* and combine them according to a *'contrast weight matrix'* to generate a '*statistic map'*. For T-tests, a '*contrast map'* along with its 'contrast *standard error map'* are also generated while for F-tests a '*contrast explained mean square map'* (i.e. the numerator of an F-statistic) is provided.

Finally, the '*inference'* activity uses a '*statistic map'* and generates an '*excursion set map'* given a '*height threshold'* and an '*extent threshold'*. The '*peak definition criteria'* and '*cluster definition criteria'* entities, used by '*inference'*, provide 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 without affecting the correction for multiple comparisons. The '*inference'* activity also generates the '*search space mask 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 in the results of a neuroimaging study.

A *'neuroimaging analysis software'* agent represents the software package used to compute the analysis. This agent is associated with all activities within the bundle.

Provenance of the *'NIDM-Results bundle'* is also recorded: the bundle was generated by a *'NIDM-Results Export'* activity which was performed by a *'NIDM-Results Exporter'* software agent corresponding to the software used to create the NIDM-Results document (e.g. FSL's Python scripts, named *'nidmfsl'* or SPM's exporter named *'spm_results_nidm'*). The bundle is associated with a version number corresponding to the version of NIDM-Results model in use.

Each activity, entity and agent has a number of predefined attributes. For instance, the list of attributes of an '*error model'* entity is provided in Fig. 1.

Updates

Each release of NIDM-Results is associated with a version number. Comments on the current version as well as suggestions of extension can be provided on the GitHub nidm repository[16]. Each extension or proposition of update will be reviewed and discussed with the members of the INCF NIDASH task force.

Implementation

SPM12 natively supports export of its results into a NIDM-Results pack, either by the use of a contextual menu in the results table or non-interactively via the batch interface as illustrated in Fig. 3. Export of FEAT results from FSL into a NIDM-results pack can be performed using the Python module nidmfsl[17], as also illustrated in Fig. 3. nidmfsl was integrated in NeuroVault and as a plugin[18] of the CBRAIN web platform for high-performance computing (Sherif et al. 2014). As a result, any FSL FEAT analysis uploaded to NeuroVault or performed in CBRAIN can be exported as a NIDM-Results pack. NeuroVault also accepts NIDM-Results packs as a mean to upload new data to a collection.

[16] https://github.com/incf-nidash/nidm
[17] https://pypi.python.org/pypi/nidmfsl
[18] https://github.com/glatard/cbrain-plugins-nidm

Publically available NIDM-Results packs

A set of 244 NIDM-Results packs has been made publically available on NeuroVault[19]. Those packs describe the results of fMRI analyses performed at the subject (232 packs) and group (12 packs) levels on six datasets downloaded from OpenfMRI[20] (R. A. Poldrack et al. 2013) (ds000005 1.1.0, ds000008 1.1.1, ds000011 unrevisioned, ds000052 unrevisioned, ds000107 unrevisioned, ds000114 unrevisioned).

Examples of usage

Meta-analysis

From 21 pain studies (10 analysed in SPM and 11 in FSL) represented in NIDM-Results we performed group coordinate-based and image-based meta-analyses contrasting the effect of pain. The data and Python script used to perform these meta-analyses are available on NeuroVault[21] and GitHub[22] respectively.

Fig. 4 provides a schematic overview of the different steps involved to compute the meta-analyses. A set of NIDM-Results packs is queried in order to retrieve the information of interest that is then combined to perform the meta-analysis.

The image-based meta-analysis was performed by combining the contrast estimate maps, along with their standard error, in a third-level mixed-effects general linear model (GLM). Each NIDM-Results pack 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 query used to extract these data is displayed in Fig. 5. The name of the corresponding contrast was associated to each map to allow for the selection of the appropriate contrast. The neuroimaging software package 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 ($p<0.05$ FWE cluster-wise with a cluster-forming threshold of $p<0.001$ uncorrected) found by the pain meta-analysis are displayed in Fig. 6. Results are also available on NeuroVault[23].

The coordinate-based meta-analysis was performed using a Multilevel Kernel Density Analysis (MKDA) (Kober et al. 2008). Each NIDM-Results pack was queried to retrieve the coordinates of the local maxima, the reference space in use and the number of subjects for each contrast. Areas of significant activation ($p<0.05$ FWE cluster-wise with a cluster-forming threshold of $p<0.001$ uncorrected) found by the pain meta-analysis are displayed in Fig. 6.

In line with previous results from the literature (Salimi-khorshidi et al. 2009), the detections for the coordinate-based and image-based meta-analysis appear consistent with a lower sensibility of the coordinate-based meta-analysis.

Reporting of neuroimaging results

Table 2 provides a mapping between the guidelines provided in (R. A. Poldrack et al. 2008) to report neuroimaging results and the fields available in a NIDM-Results serialisation. NIDM-Results cover all elements from the "Statistical modelling" checklist that could be automatically retrieved within the neuroimaging software package.

[19] http://neurovault.org/collections/1435/
[20] https://openfmri.org/
[21] http://neurovault.org/collections/NJXQPEEK/
[22] https://github.com/incf-nidash/nidmresults-paper/
[23] http://neurovault.org/collections/1432/

Examples of reports generated from a NIDM-Results export of group and single-subject analyses performed in SPM and FSL are provided in Fig. 7. The data and Python script used to generate those report are available on NeuroVault[19] and GitHub[22] respectively.

Discussion

Data sharing in the neuroimaging community is still restrained by a number of psychological and ethical factors that are beyond the scope of the current paper (see (R. A. Poldrack and Gorgolewski 2014; Poline et al. 2012) for a review). Those will have to be addressed in order for data sharing to become common practice in the neuroimaging community. In an effort to address the technological barriers that make data sharing challenging, here we have proposed a solution to share neuroimaging results of mass univariate analyses.

As a first step to provide machine-readable metadata, we restricted our scope to information that was automatically extractable and attributes that were crucial for meta-analysis (e.g. number of subjects). This limited the amount of information that could be represented. For instance, the description of the paradigm was limited to the design matrix and 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[24] (R. a. Poldrack et al. 2011) and CogPO[25] (Turner and Laird 2012)), description of fMRI paradigms is still a topic of active research. Some level of manual interaction to select contrasts of interest is therefore needed to compute a meta-analysis based on NIDM-Results packs. 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. Similarly, NIDM-Results could be extended to match emerging best practices (such as (Nichols et al. 2016)).

NIDM-Results currently focuses on the representation of parametric mass-univariate analyses. Thanks to the intrinsic extensibility of RDF models, variants could be proposed to broaden its scope. For example, an extension for non-parametric statistics is under discussion[26]. Mass-univariate results, as the most well established approach for fMRI analysis, was an obvious choice to start a cross-software modelling effort. But neuroimaging cannot be limited to mass-univariate analyses and future work will focus on providing extensions for other types of analyses (e.g. analysis of resting state fMRI).

We based our modelling effort on PROV, a specification endorsed by the W3C, to model provenance on the web. Other efforts have been proposed to model provenance including families of ontologies like the OBO foundry (B. Smith et al. 2007) or DOLCE (Masolo et al. 2003). We chose PROV as it is lightweight, focused only on provenance, and is easily extensible to provide domain-specific knowledge.

NIDM-Results is based on RDF and semantic web technologies. While a number of ontologies have been developed in relation with neuroimaging (e.g. Cognitive Atlas (Poldrack et al. 2011), CogPO (Turner and Laird 2012), OntoNeurolog (Temal et al. 2008)), the use of controlled vocabularies and of linked data is not yet common practice in our community. As more and more data become available online and as standardisation effort like the RII develops, we believe that these technologies will become more widespread. RDF was chosen as a basis for NIDM for the expressivity of its graph-based structures, the possibility to form intricate queries across datasets (D. B. Keator et al. 2013), as well as for the extensibility of the created data models and the possibility to interconnect across knowledge domains (cf. (Nolan Nichols et al. 2016) for a review).

[24] http://www.cognitiveatlas.org/
[25] http://www.cogpo.org/
[26] https://github.com/incf-nidash/nidm/pull/233

One limitation of NIDM-Results is that only limited provenance is represented. For instance, computational environment, which has been shown to be source of undesired variability in neuroimaging results (Glatard et al. 2015), is not part of our model. NIDM-Results is part of a broader effort (NIDM) that aims at representing different levels of provenance in neuroimaging experiments. 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.

As for the definition of any new model, 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 from the community is welcome and can be submitted as issues in our GitHub repository[27]. We also made a strong commitment to make implementations available. Taking advantage of the fact that most functional MRI studies are performed using a limited number of software packages (> 75% for SPM or FSL, > 90% for SPM, FSL or AFNI according to (Carp 2012)), we developed implementations for SPM and FSL, and are currently working with AFNI developers to further extend the coverage of NIDM-Results export.

While we have focused our implementation efforts on the generation of NIDM-Results packs, the development of applications processing NIDM-Results is also crucial, to serve as incentives for neuroimaging users. As an example, 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. A wider ecosystem is under development (including a standalone viewer).

We believe NIDM-Results is an essential tool for the future of transparent, reproducible science using neuroimaging. If all research publications were accompanied by such a machine-readable description of the experiment, debates on the exact methodology used would be compressed or eliminated, and any replication efforts greatly facilitated.

Material and methods

Process

Since August 2013 the model was developed through weekly teleconferences and eight focused workshops during which the team of experts iteratively defined the terminology, seeking to ensure that the output of AFNI, FSL and SPM could be represented in this framework. Furthermore, a separate meeting was organised with each of the development teams of 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[28] and on GitHub under the incf-nidash organization[29].

Scope of the model

NIDM-Results focuses on mass-univariate models based on a General Linear Model (at the subject or group level). To facilitate adoption, we restricted the scope of NIDM-Results to metadata that could be automatically extracted with limited user input, motivated by the specific metadata that is crucial for the application of meta-analysis. 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 (R. A. Poldrack et al. 2008), and (3), in an effort to ensure continuity with

[27] https://github.com/incf-nidash/nidm/issues/new
[28] https://drive.google.com/open?id=0B-BLof5_SOh8bWR3UDE4WTdELXM
[29] https://github.com/incf-nidash/nidm

current practice, we also considered the elements displayed as part of results reporting in different neuroimaging software (e.g. peaks, clusters). When an item, essential for image-based meta-analysis, was not produced as part of the standard analysis (e.g. the contrast standard error map in SPM) we included it in the model and depend on the exporters to generate it from existing data.

Term re-use and definitions

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, NeuroLex for neuroscience terms, RRID for tools and also, to a lesser extent, Dublin Core, the NEPOMUK file ontology and the Cryptographic Hash Functions vocabulary. Namespaces of the re-used ontologies are provided in Table 3. More details on the re-used vocabularies are provided below.

PROV

The W3C specification PROV defines three types of objects: an *Activity* represents a process that was performed on some data (e.g. a voxel-wise inference) and occurred over a fixed period of time; an *Agent* represents someone (human, organization, machine…) that takes responsibility for an activity (e.g. the 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 voxel-wise inference *Activity used* a NIfTI image *Entity*; a voxel-wise inference *Activity was associated with* the SPM *Agent* and another NIfTI image *Entity was generated by* the segmentation *Activity*). NIDM-Results terms were defined as specialisations of PROV terms.

STATO

GLM analyses of fMRI data rely on well-known statistical constructs (e.g. one-sample T-test, two-sample T-test, F-tests, ANOVA, inference, ordinary least squares estimation, etc.). The general-purpose STATistics Ontology (STATO) is built on the top of the OBO foundry and aims to provide a set of terms describing statistics. We re-used statistics terms available in STATO (e.g. obo:'t-statistic') and when we could not find an appropriate statistical term, we engaged with STATO developers through GitHub issues[30] to propose new terms (e.g. "residual mean squares" discussed in issue 35[31]).

NeuroLex and RRID

Much work has been done in the neuroimaging community to provide controlled vocabularies and ontologies defining neuroimaging concepts. NeuroLex[32] (Larson and Martone 2013) provides a common platform that gathers terms from different sources (including previous vocabularies developed by NIF, BIRN…). Interestingly, Neurolex was part of the recent Resource Identification Initiative[33] (RII) (Bandrowski and Martone 2016) that publicized the use of those identifiers (e.g. “RRID:SCR_007037” for SPM[34]) 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. We re-used the available RRIDs describing neuroimaging software packages.

Dublin Core, NEPOMUK file ontology and the Cryptographic Hash Functions vocabulary

Many vocabularies and ontologies have terms available to describe files. We chose to rely on the widely adopted DUBLIN core terminology. Additionally, we used the “fileName” term from the

[30] https://github.com/ISA-tools/stato/issues?utf8=%E2%9C%93&q=is%3Aissue+nidm+
[31] https://github.com/ISA-tools/stato/issues/35
[32] http://neurolex.org/
[33] https://scicrunch.org/resources
[34] https://scicrunch.org/resolver/nif-0000-00343

NEPOMUK file ontology and the SHA-256 term from the Cryptographic Hash Functions vocabulary.

New terms

When no term was found to describe a given neuroimaging concept of interest, we created a new term and carefully crafted a definition or engaged with the relevant ontology maintainers (e.g. STATO) to propose a new definition. All new 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).

Examples of usage

Meta-analysis

Results from 21 pain studies previously analysed with FSL were made available to us. The second-level analyses were recomputed with SPM for 10 of those studies in order to obtain a dataset of NIDM-Results packs coming from mixed software packages. We computed a one-sample meta-analysis contrasting the effect of "pain" and compared the results of coordinate-based and image-based meta-analyses.

The MKDA toolbox[35] was used to perform the coordinate-based meta-analysis. The nidmresults Python toolbox was used to generate the csv file required as input for the analyses.

FSL's FLAME 1 (S. Smith et al. 2001) was used to compute the image-based meta-analysis with the gold standard approach (3rd level mixed-effects general linear model). Difference in data scaling between software packages were compensated 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 using the RDFlib library[36] to query the document. The paragraph that was generated could, for instance, be used as part of the method section in a research paper.

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. We also gratefully acknowledge the Tracey group at FMRIB for sharing their pain datasets used in the meta-analysis.

Funding

The INCF supported and organised the task force meetings in which the model was discussed. AB, CM and TEN were supported by the Wellcome Trust [100309/Z/12/Z]. TA was supported by the Medical Research Council (United Kingdom) [MC-A060-53114]. BNN was supported by NIH grants [AA012388, AA021697, AA021697-04S1]. SG was partially supported by NIH grants [1R01EB020740-01A1, 1P41EB019936-01A1]. JBP was partially supported by a NIH-NIBIB grant [P41-EB019936], the Laura and John Arnold Foundation and by a NIH-NIDA grant [U24-038653]. DBK was supported by the Function Biomedical Informatics Re-search Network (NIH 1 [U24 U24 RR021992]), the BIRN Coordinating Center (https://www.birncommunity.org; NIH 1 [U24 RR025736-01]) and the Conte Center on Brain Programming in Adolescent Vulnerabilities [1P50MH096889-01A1]. GC and RR were supported by Supported by the NIMH and NINDS

[35] http://wagerlab.colorado.edu/tools
[36] https://github.com/RDFLib/rdflib/

Intramural Research Programs (ZICMH002888) of the NIH/HHS, USA. CG was sponsored by the Laura and John Arnold Foundation. SD and TG were supported by the Irving Ludmer Family Foundation and the Ludmer Centre for Neuroinformatics and Mental Health.

References

Bandrowski, Anita E., and Maryann E. Martone. 2016. "RRIDs: A Simple Step toward Improving Reproducibility through Rigor and Transparency of Experimental Methods." *Neuron* 90 (3): 434–36. doi:10.1016/j.neuron.2016.04.030.

Button, Katherine S., John P. A. Ioannidis, Claire Mokrysz, Brian A. Nosek, Jonathan Flint, Emma S. J. Robinson, and Marcus R. Munafò. 2013. "Power Failure: Why Small Sample Size Undermines the Reliability of Neuroscience." *Nature Reviews. Neuroscience* 14 (5). Nature Publishing Group: 365–76. doi:10.1038/nrn3475.

Carp, Joshua. 2012. "The Secret Lives of Experiments: Methods Reporting in the fMRI Literature." *NeuroImage* 63 (1). Elsevier Inc.: 289–300. doi:10.1016/j.neuroimage.2012.07.004.

———. 2013. "Better Living through Transparency: Improving the Reproducibility of fMRI Results through Comprehensive Methods Reporting." *Cognitive, Affective & Behavioral Neuroscience* 13 (3): 660–66. doi:10.3758/s13415-013-0188-0.

Costafreda, Sergi G., Anthony S. David, and Michael J. Brammer. 2009. "A Parametric Approach to Voxel-Based Meta-Analysis." *NeuroImage* 46 (1): 115–22. doi:10.1016/j.neuroimage.2009.01.031.

Cox, R. W. 1996. "AFNI: Software for Analysis and Visualization of Functional Magnetic Resonance Neuroimages." *Computers and Biomedical Research, an International Journal*, 162–73. http://www.sciencedirect.com/science/article/pii/S0010480996900142.

Eickhoff, Simon B., Danilo Bzdok, Angela R. Laird, Florian Kurth, and Peter T. Fox. 2012. "Activation Likelihood Estimation Meta-Analysis Revisited." *NeuroImage* 59 (3): 2349–61. doi:10.1016/j.neuroimage.2011.09.017.

Gadde, Syam, Nicole Aucoin, Jeffrey S. Grethe, David B. Keator, Daniel S. Marcus, and Steve Pieper. 2012. "XCEDE: An Extensible Schema for Biomedical Data." *Neuroinformatics* 10 (1): 19–32. doi:10.1007/s12021-011-9119-9.

Glatard, Tristan, Lindsay B. Lewis, Rafael Ferreira da Silva, Reza Adalat, Natacha Beck, Claude Lepage, Pierre Rioux, et al. 2015. "Reproducibility of Neuroimaging Analyses across Operating Systems." *Frontiers in Neuroinformatics* 9 (April): 12. doi:10.3389/fninf.2015.00012.

Gorgolewski, Krzysztof, Christopher D. Burns, Cindee Madison, Dav Clark, Yaroslav O. Halchenko, Michael L. Waskom, and Satrajit S. Ghosh. 2011. "Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python." *Frontiers in Neuroinformatics* 5 (August): 13. doi:10.3389/fninf.2011.00013.

Gorgolewski, Krzysztof J., Gael Varoquaux, Gabriel Rivera, Yannick Schwarz, Satrajit S. Ghosh, Camille Maumet, Vanessa V. Sochat, et al. 2015. "NeuroVault.org: A Web-Based Repository for Collecting and Sharing Unthresholded Statistical Maps of the Human Brain." *Frontiers in Neuroinformatics* 9 (April): 8. doi:10.3389/fninf.2015.00008.

Huynh, Trung Dong, Paul Groth, and Stephan Zednik. 2013. "PROV Implementation Report." Edited by Trung Dong Huynh, Paul Groth, and Stephan Zednik, April. World Wide Web Consortium. http://eprints.soton.ac.uk/358440/.

Inglis, Ben. 2015. "A Checklist for fMRI Acquisition Methods Reporting in the Literature." *The Winnower*. doi:10.15200/winn.143191.17127.

Jenkinson, Mark, Christian F. Beckmann, Timothy E. J. Behrens, Mark W. Woolrich, and

Stephen M. Smith. 2012. “FSL.” *NeuroImage* 62 (2): 782–90. doi:10.1016/j.neuroimage.2011.09.015.
Keator, David B., Dingying Wei, Syam Gadde, Jeremy Bockholt, Jeffrey S. Grethe, Daniel Marcus, Nicole Aucoin, and Ibrahim B. Ozyurt. 2009. “Derived Data Storage and Exchange Workflow for Large-Scale Neuroimaging Analyses on the BIRN Grid.” *Frontiers in Neuroinformatics* 3 (September): 30. doi:10.3389/neuro.11.030.2009.
Keator, D. B., Syam Gadde, J. S. Grethe, D. V. Taylor, and S. G. Potkin. 2006. “A General XML Schema and SPM Toolbox for Storage of Neuro-Imaging Results and Anatomical Labels.” *Neuroinformatics* 00: 199–211. doi:10.1385/NI.
Keator, D. B., K. Helmer, J. Steffener, J. A. Turner, T. G. M. Van Erp, S. Gadde, N. Ashish, G. A. Burns, and B. N. Nichols. 2013. “Towards Structured Sharing of Raw and Derived Neuroimaging Data across Existing Resources.” *NeuroImage* 82 (November). Elsevier B.V.: 647–61. doi:10.1016/j.neuroimage.2013.05.094.
Kober, Hedy, Lisa Feldman Barrett, Josh Joseph, Eliza Bliss-Moreau, Kristen Lindquist, and Tor D. Wager. 2008. “Functional Grouping and Cortical-Subcortical Interactions in Emotion: A Meta-Analysis of Neuroimaging Studies.” *NeuroImage* 42 (2): 998–1031. doi:10.1016/j.neuroimage.2008.03.059.
Laird, Angela R., Jack L. Lancaster, and Peter T. Fox. 2005. “BrainMap: The Social Evolution of a Human Brain Mapping Database.” *Neuroinformatics* 3 (1): 65–78. http://www.ncbi.nlm.nih.gov/pubmed/15897617.
Larson, Stephen D., and Maryann E. Martone. 2013. “NeuroLex.org: An Online Framework for Neuroscience Knowledge.” *Frontiers in Neuroinformatics* 7 (August): 18–18. doi:10.3389/fninf.2013.00018.
Masolo, Claudio, Stefano Borgo, Aldo Gangemini, Nicola Guarino, Alessandro Oltramari, and L. Schneider. 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)*.
Moreau, Luc, and Paolo Missier. 2013. “PROV-DM: The PROV Data Model.” Edited by Luc Moreau and Paolo Missier. World Wide Web Consortium. http://eprints.soton.ac.uk/356851/.
Nichols, Thomas E., Samir Das, Simon B. Eickhoff, Alan C. Evans, Tristan Glatard, Michael Hanke, Nikolaus Kriegeskorte, et al. 2016. “Best Practices in Data Analysis and Sharing in Neuroimaging Using MRI.” doi:10.1101/054262.
Nolan Nichols, B., Satrajit S. Ghosh, Tibor Auer, Thomas J. Grabowskith, Camille Maumet, David Keator, Kilian Pohl, and Jean-Baptiste Poline. 2016. “Building a Web of Linked Data Resources to Advance Neuroscience Research.” *bioRxiv*. doi:10.1101/053934.
Penny, William D., Karl J. Friston, John T. Ashburner, Stefan J. Kiebel, and Thomas E. Nichols. 2011. *Statistical Parametric Mapping: The Analysis of Functional Brain Images: The Analysis of Functional Brain Images*. Academic press. https://books.google.com/books?hl=en&lr=&id=G_qdEsDlkp0C&oi=fnd&pg=PP1&dq=Statistical%2BParametric%2BMapping%2BThe%2BAnalysis%2Bof%2BFunctional%2BBrain%2BImages&ots=Xl1IDBS6YH&sig=vIuutu87of5R0yNMhZny-y4Mxc0.
Poldrack, Russell a., Aniket Kittur, Donald Kalar, Eric Miller, Christian Seppa, Yolanda Gil, D. Stott Parker, Fred W. Sabb, and Robert M. Bilder. 2011. “The Cognitive Atlas: Toward a Knowledge Foundation for Cognitive Neuroscience.” *Frontiers in Neuroinformatics* 5 (September): 17–17. doi:10.3389/fninf.2011.00017.
Poldrack, Russell A., Deanna M. Barch, Jason P. Mitchell, Tor D. Wager, Anthony D. Wagner, Joseph T. Devlin, Chad Cumba, Oluwasanmi Koyejo, and Michael P. Milham. 2013. “Toward Open Sharing of Task-Based fMRI Data: The OpenfMRI

Project." *Frontiers in Neuroinformatics* 7 (July): 12–12. doi:10.3389/fninf.2013.00012.
Poldrack, Russell A., Paul C. Fletcher, Richard N. Henson, Keith J. Worsley, Matthew Brett, and Thomas E. Nichols. 2008. "Guidelines for Reporting an fMRI Study." *NeuroImage* 40 (2): 409–14. doi:10.1016/j.neuroimage.2007.11.048.
Poldrack, Russell A., and Krzysztof J. Gorgolewski. 2014. "Making Big Data Open: Data Sharing in Neuroimaging." *Nature Neuroscience* 17 (11): 1510–17. doi:10.1038/nn.3818.
Poline, Jean-Baptiste, Janis L. Breeze, Satrajit Ghosh, Krzysztof Gorgolewski, Yaroslav O. Halchenko, Michael Hanke, Christian Haselgrove, et al. 2012. "Data Sharing in Neuroimaging Research." *Frontiers in Neuroinformatics* 6 (April): 9–9. doi:10.3389/fninf.2012.00009.
Rex, David E., Jeffrey Q. Ma, and Arthur W. Toga. 2003. "The LONI Pipeline Processing Environment." *NeuroImage* 19 (3): 1033–48. doi:10.1016/S1053-8119(03)00185-X.
Salimi-khorshidi, Gholamreza, Stephen M. Smith, John R. Keltner, Tor D. Wager, and Thomas E. Nichols. 2009. "Meta-Analysis of Neuroimaging Data: A Comparison of Image-Based and Coordinate-Based Pooling of Studies." *NeuroImage* 45 (3). Elsevier B.V.: 810–23. doi:10.1016/j.neuroimage.2008.12.039.
Sherif, Tarek, Pierre Rioux, Marc-Etienne Rousseau, Nicolas Kassis, Natacha Beck, Reza Adalat, Samir Das, Tristan Glatard, and Alan C. Evans. 2014. "CBRAIN: A Web-Based, Distributed Computing Platform for Collaborative Neuroimaging Research." *Frontiers in Neuroinformatics* 8 (May): 54. doi:10.3389/fninf.2014.00054.
Smith, Barry, Michael Ashburner, Cornelius Rosse, Jonathan Bard, William Bug, Werner Ceusters, Louis J. Goldberg, et al. 2007. "The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration." *Nature Biotechnology* 25 (11): 1251–55. doi:10.1038/nbt1346.
Smith, Stephen, Peter R. Bannister, Christian Beckmann, Mike Brady, Stuart Clare, David Flitney, Peter Hansen, et al. 2001. "FSL: New Tools for Functional and Structural Brain Image Analysis." *NeuroImage* 13 (6). Academic Press: 249.
Temal, Lynda, Michel Dojat, Gilles Kassel, and Bernard Gibaud. 2008. "Towards an Ontology for Sharing Medical Images and Regions of Interest in Neuroimaging." *Journal of Biomedical Informatics* 41: 766–78. doi:10.1016/j.jbi.2008.03.002.
Turner, Jessica a., and Angela R. Laird. 2012. "The Cognitive Paradigm Ontology: Design and Application." *Neuroinformatics* 10: 57–66. doi:10.1007/s12021-011-9126-x.
Yarkoni, Tal, Russell A. Poldrack, Thomas E. Nichols, David C. Van Essen, Tor D. Wager, and David C. Van Essen. 2011. "Large-Scale Automated Synthesis of Human Functional Neuroimaging Data." *Nature Methods* 8 (8): 665–70. doi:10.1038/nmeth.1635.

Figures

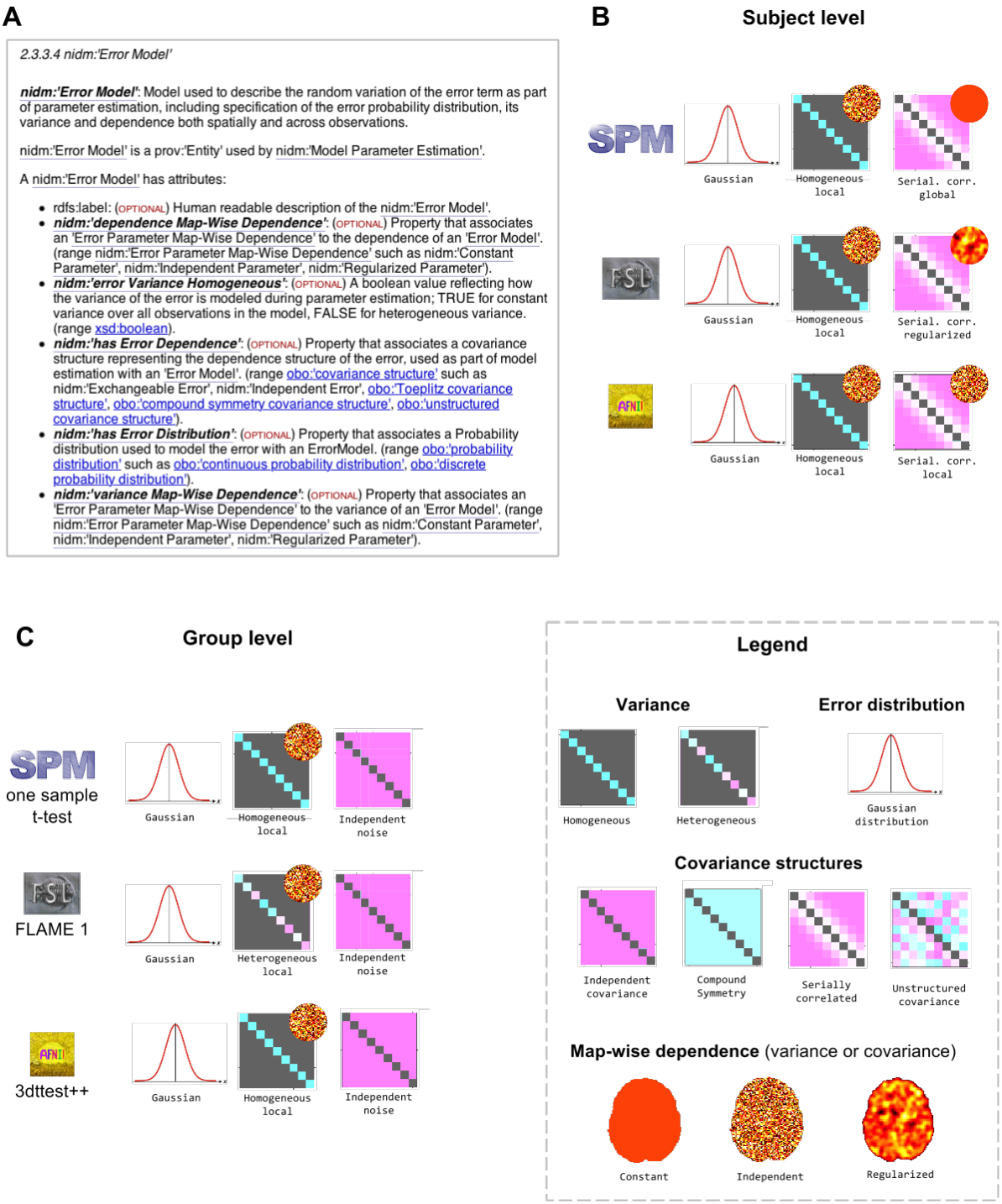

Fig. 1. Description of the error models with NIDM-Results. Excerpt of the NIDM-Results 1.3.0 specification describing a nidm:'Error Model' and its attributes (A). Examples of model implementations for subject-level (B) and group-level (C) analyses for SPM, FSL and AFNI.

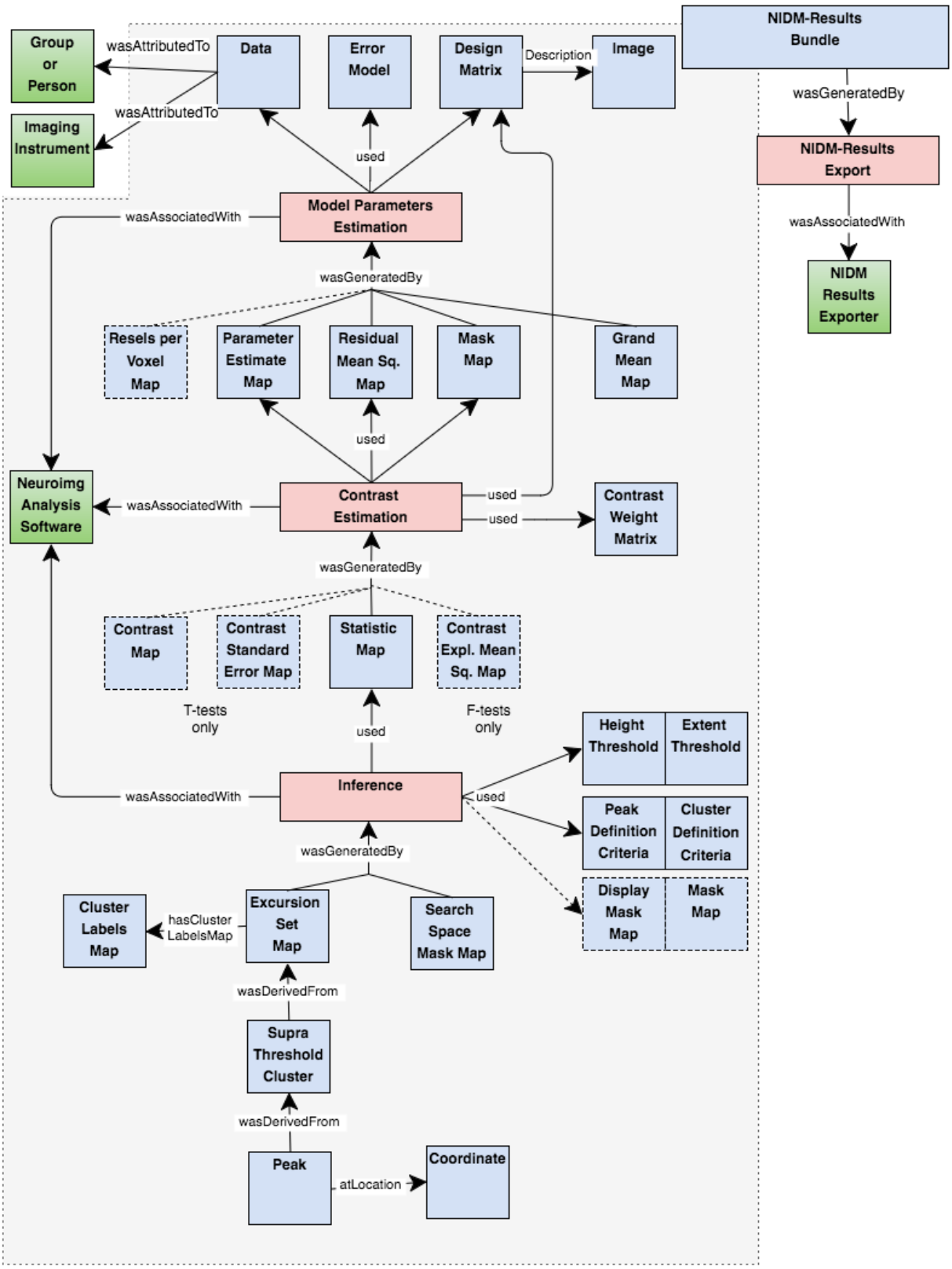

Fig. 2. NIDM-Results objects. Color-coding indicates the type as defined in PROV (blue: Entity, red: Activity, green: Agent).

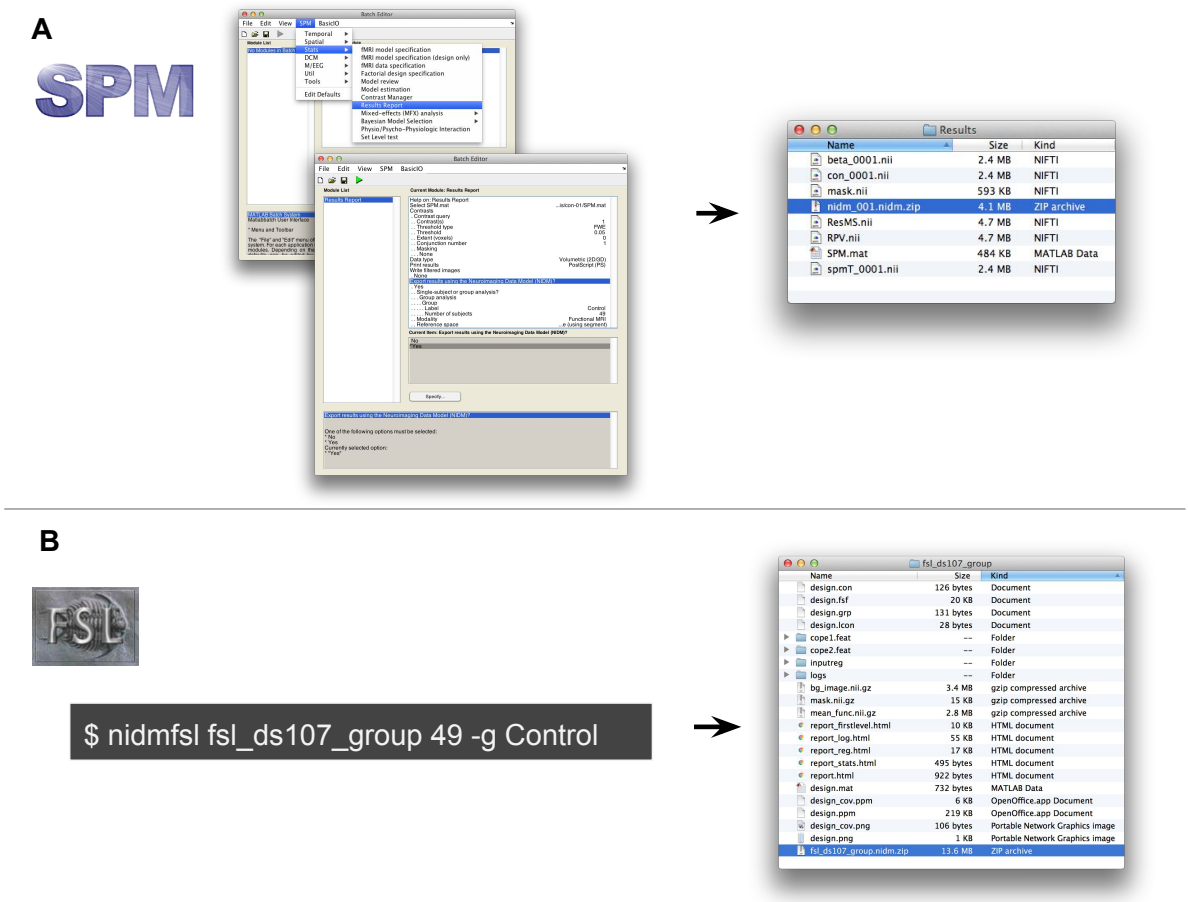

Fig. 3. NIDM-Results export in SPM12 (A) and FSL v5.0 (B).

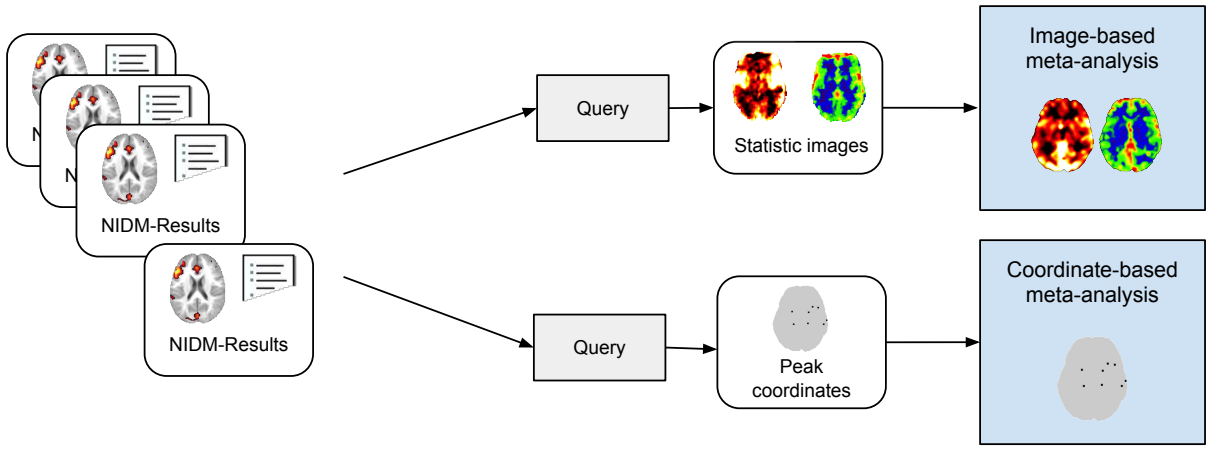

Fig. 4. Image-based and coordinate-based meta-analyses using NIDM-Results. Each NIDM-Results pack is queried to retrieve the data and metadata of interest for each type of meta-analysis. These data are then combined in a meta-analysis.

```
prefix prov: <http://www.w3.org/ns/prov#>
prefix nidm: <http://purl.org/nidash/nidm#>
prefix contrast_estimation: <http://purl.org/nidash/nidm#NIDM_0000001>
prefix contrast_map: <http://purl.org/nidash/nidm#NIDM_0000002>
prefix stderr_map: <http://purl.org/nidash/nidm#NIDM_0000013>
prefix contrast_name: <http://purl.org/nidash/nidm#NIDM_0000085>
prefix statistic_map: <http://purl.org/nidash/nidm#NIDM_0000076>
prefix mask_map: <http://purl.org/nidash/nidm#NIDM_0000054>

SELECT ?contrastName ?con_file ?std_file ?mask_file ?software
WHERE {
      ?con_id a contrast_map: ;
          contrast_name: ?contrastName ;
          prov:atLocation ?con_file ;
          prov:wasGeneratedBy ?con_est .
      ?std_id a stderr_map: ;
          prov:atLocation ?std_file ;
          prov:wasGeneratedBy ?con_est .
      ?mask_id a mask_map: ;
          prov:atLocation ?mask_file .
      ?soft_id a ?software .
      ?con_est a contrast_estimation: ;
            prov:wasAssociatedWith ?soft_id ;
            prov:used ?mask_id .
       FILTER(?software NOT IN (prov:SoftwareAgent, prov:Agent))
}
```

Fig. 5. SPARQL query to retrieve data and metadata needed for image-based meta-analysis (syntax was highlighted using CodeMirror[37])

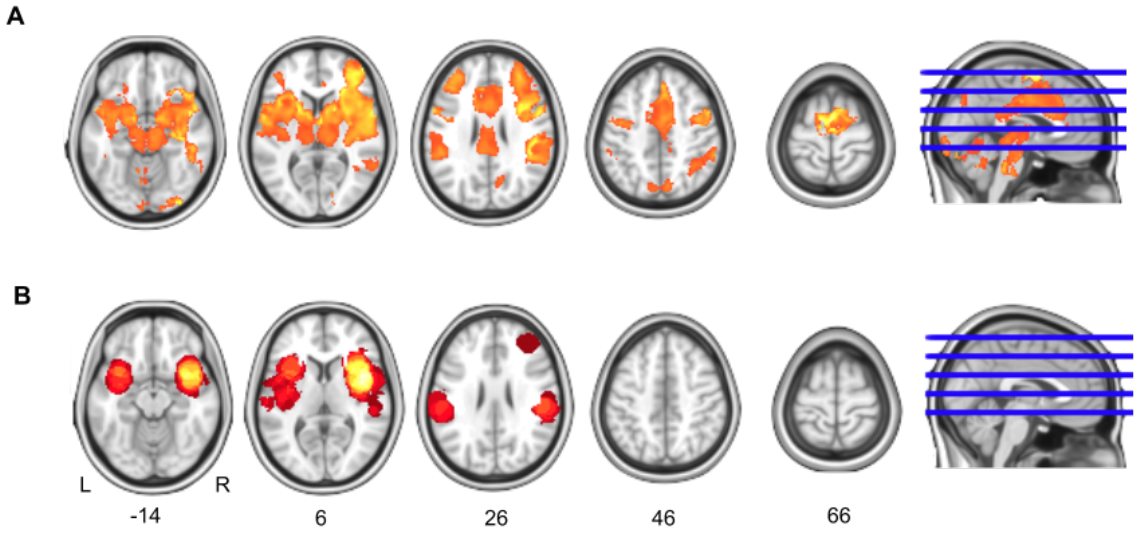

Fig. 6. One-sample meta-analysis of 21 studies investigating the effect of pain. Areas of significant activation with an FWE-corrected cluster-wise threshold p<0.05 (cluster-forming threshold p<0.001 uncorrected) for the image-based (A) and the coordinate-based (B) meta-analyses.

[37] https://codemirror.net

A FSL group analysis

Group-level analysis was performed with FSL (version 5.0.x). A linear regression was computed at each voxel, using weighted least squares (assuming unequal variances) with a local variance estimate.
Cluster-wise inference was performed with correction for multiple comparisons using a threshold P <= 0.050 (FWER adjusted) with a cluster defining threshold Z-statistic >= 2.300. The search volume was 1522 cm^3 (190327 voxels).

B SPM group analysis

Group-level analysis was performed with SPM (version 12.6685). A linear regression was computed at each voxel, using ordinary least squares (assuming equal variances) with a local variance estimate.
Voxel-wise inference was performed with correction for multiple comparisons using a threshold P <= 0.050 (FWER adjusted). The search volume was 949 cm^3 (118626 voxels).

C FSL single-subject analysis

Subject-level analysis was performed with FSL (version 5.0.x). A linear regression was computed at each voxel, using generalized least squares (assuming equal variances) with a local variance estimate and a spatially regularized Toeplitz covariance structure. Drift was fit with a gaussian running line drift model (60.0s FWHM).
Voxel-wise inference was performed using a threshold P <= 0.001 (Uncorrected). The search volume was 1539 cm^3 (57029 voxels).

D SPM single-subject analysis

Subject-level analysis was performed with SPM (version 12.6685). A linear regression was computed at each voxel, using generalized least squares (assuming equal variances) with a local variance estimate and a global Toeplitz covariance structure. Drift was fit with a discrete cosine transform basis drift model (128.0s cut-off).
Voxel-wise inference was performed using a threshold P <= 0.001 (Uncorrected). The search volume was 1791 cm^3 (223883 voxels).

Fig. 7. Examples of reports generated from NIDM-Results packs for group (A, B) and single-subject (C, D) analyses performed in FSL (A, C) and SPM (B, D).

Tables

PROV type	Term	Qualified name
Entity	NIDM-Results bundle	nidm:NIDM_0000027
	Bundle	prov:Bundle
	Design Matrix	nidm:NIDM_0000019
	Error Model	nidm:NIDM_0000023
	Data	nidm:NIDM_0000169
	Parameter Estimate Map(s)	nidm:NIDM_0000061
	Mask Map	nidm:NIDM_0000054
	Residual Mean Squares Map	nidm:NIDM_0000066
	Resels Per Voxel Map	nidm:NIDM_0000144
	Grand Mean Map	nidm:NIDM_0000033
	contrast weight matrix	obo:STATO_0000323
	Statistic Map	nidm:NIDM_0000076
	Contrast Map	nidm:NIDM_0000002
	Contrast Standard Error Map	nidm:NIDM_0000013
	Contrast Explained Mean Square Map	nidm:NIDM_0000163
	Excursion Set Map	nidm:NIDM_0000025
	Height Threshold	nidm:NIDM_0000034
	Extent Threshold	nidm:NIDM_0000026
	Peak Definition Criteria	nidm:NIDM_0000063
	Cluster Definition Criteria	nidm:NIDM_0000007
	Display Mask Map	nidm:NIDM_0000020
	Search Space Mask Map	nidm:NIDM_0000068
	Supra-Threshold Cluster(s)	nidm:NIDM_0000070
	Peak(s)	nidm:NIDM_0000062
Activity	Model Parameter Estimation	nidm:NIDM_0000056
	Contrast Estimation	nidm:NIDM_0000001

	Inference	nidm:NIDM_0000049
	Conjunction Inference	nidm:NIDM_0000011
	NIDM-Results Export	nidm:NIDM_0000166
Agent	Neuroimaging Analysis Software	nidm:NIDM_0000164
	Person	prov:Person
	study group population	obo:STATO_0000193
	Imaging Instrument	nif:birnlex_2094
	NIDM-Results Exporter	nidm:NIDM_0000165
	nidmfsl	nidm:NIDM_0000167
	spm_results_nidm	nidm:NIDM_0000168

Table 1. PROV type, label and identifier of the NIDM-Results terms mentioned in single quotes in this manuscript.

(**Table 2** is available below table 3).

Vocabulary/Ontology	**URI**	**Prefix**
PROV	http://www.w3.org/ns/prov#	prov
STATO	http://purl.obolibrary.org/obo/	obo
NeuroLex	http://uri.neuinfo.org/nif/nifstd/	nlx
RRID	http://scicrunch.org/resolver/	rrid
Dublin Core types	http://purl.org/dc/dcmitype/	dctype
Dublin Core elements	http://purl.org/dc/elements/1.1/	dc
Dublin Core terms	http://purl.org/dc/terms/	dct
Cryptographic Hash Functions	http://id.loc.gov/vocabulary/preservation/cryptographicHashFunctions#	crypto
NEPOMUK file ontology	http://www.semanticdesktop.org/ontologies/2007/03/22/nfo#	nfo
NIDM	http://purl.org/nidash/nidm#	nidm
FSL	http://purl.org/nidash/fsl#	fsl
SPM	http://purl.org/nidash/spm#	spm

Table 3. Prefixes of the vocabularies used in NIDM-Results.

Checklist from (Poldrack et al. 2008)	NIDM-Results representation	Example (turtle)
Intra-subject fMRI modeling info		
Estimation method	Attribute **nidm:'with Estimation Method'** of the **nidm:'Model Parameters Estimation'** activity. Possible values include: • obo:'ordinary least squares estimation' for ordinary least squares, • obo:'generalized least squares estimation' for generalized least squares or; • obo:'weighted least squares estimation' for weighted least squares.	EXAMPLE: Ordinary Least Squares Estimation @prefix nidm_ModelParametersEstimation: <http://purl.org/nidash/nidm#NIDM_0000056> . @prefix nidm_withEstimationMethod: <http://purl.org/nidash/nidm#NIDM_0000134> . @prefix obo_ordinaryleastsquaresestimation: <http://purl.obolibrary.org/obo/STATO_0000370> . niiri:model_pe_id prov:used niiri:error_model_id ; a prov:Activity , nidm_ModelParametersEstimation: ; rdfs:label "Model parameters estimation" ; **nidm_withEstimationMethod: obo_ordinaryleastsquaresestimation: .**
Hemodynamic response function	Attribute **'has HRF Basis'** of a **'Design Matrix'** entity. Possible values include: • spm:'SPM's Canonical HRF' for SPM's canonical hemodynamic response function (default in SPM). • fsl:'FSL's Gamma Difference HRF' for FSL's • nidm:'Finite Impulse Response Basis Set'	EXAMPLE: HRF: SPM's Informed Basis Set @prefix nidm_DesignMatrix: <http://purl.org/nidash/nidm#NIDM_0000019> . @prefix nidm_regressorNames: <http://purl.org/nidash/nidm#NIDM_0000021> . @prefix nidm_hasHRFBasis: <http://purl.org/nidash/nidm#NIDM_0000102> . @prefix nidm_hasDriftModel: <http://purl.org/nidash/nidm#NIDM_0000088> . @prefix spm_SPMsCanonicalHRF: <http://purl.org/nidash/spm#SPM_0000004> . @prefix spm_SPMsTemporalDerivative: <http://purl.org/nidash/spm#SPM_0000006> . @prefix spm_SPMsDispersionDerivative: <http://purl.org/nidash/spm#SPM_0000003> . niiri:first_level_design_matrix_id a prov:Entity , nidm_DesignMatrix: ; rdfs:label "First-Level Design Matrix" ; prov:atLocation "DesignMatrix.csv"^^xsd:anyURI ; dct:format "text/csv"^^xsd:string ; nfo:fileName "DesignMatrix.csv"^^xsd:string ; dc:description niiri:design_matrix_png_id ; nidm_regressorNames: "[\"Sn(1) active*bf(1)\",\"Sn(1) constant\"]"^^xsd:string ; nidm_hasDriftModel: niiri:drift_model_id ; **nidm_hasHRFBasis: spm_SPMsCanonicalHRF: ;** **nidm_hasHRFBasis: spm_SPMsTemporalDerivative: ;** **nidm_hasHRFBasis: spm_SPMsDispersionDerivative: .**
Drift modeling/high-pass filtering	Attribute **'has Drift Model'** of a **'Design Matrix'** entity. Possible values include: • fsl:'Gaussian Running Line Drift Model' for a Gaussian-weighted running line smoother • spm:'DCT Drift Model' for Discrete Cosine Transform basis.	EXAMPLE: FSL's Gaussian Running Line Drift Model @prefix fsl_GaussianRunningLineDriftModel: <http://purl.org/nidash/fsl#FSL_0000002> . @prefix fsl_driftCutoffPeriod: <http://purl.org/nidash/fsl#FSL_0000004> . **niiri:drift_model_id a prov:Entity , fsl_GaussianRunningLineDriftModel: ;** rdfs:label "FSL's Gaussian Running Line Drift Model" ; fsl_driftCutoffPeriod: "2"^^xsd:float .
Autocorrelation model		

Model type	Attribute **'has Error Dependence'** of an **'Error Model'** entity. Possible values include: • obo:'Toeplitz covariance structure' for serially correlated error • obo:'unstructured covariance structure' for arbitrary autocorrelation function	EXAMPLE: Error Model: SPM group analysis with non sphericity @prefix nidm_ErrorModel: <http://purl.org/nidash/nidm#NIDM_0000023> . @prefix nidm_hasErrorDistribution: <http://purl.org/nidash/nidm#NIDM_0000101> . @prefix nidm_errorVarianceHomogeneous: <http://purl.org/nidash/nidm#NIDM_0000094> . @prefix nidm_varianceMapWiseDependence: <http://purl.org/nidash/nidm#NIDM_0000126> . @prefix nidm_hasErrorDependence: <http://purl.org/nidash/nidm#NIDM_0000100> . @prefix nidm_dependenceMapWiseDependence: <http://purl.org/nidash/nidm#NIDM_0000089> . @prefix nidm_IndependentParameter: <http://purl.org/nidash/nidm#NIDM_0000073> . @prefix nidm_ConstantParameter: <http://purl.org/nidash/nidm#NIDM_0000072> . @prefix obo_normaldistribution: <http://purl.obolibrary.org/obo/STATO_0000227> . @prefix obo_unstructuredcovariancestructure: <http://purl.obolibrary.org/obo/STATO_0000405> . niiri:error_model_id a prov:Entity , nidm_ErrorModel: ; nidm_hasErrorDistribution: obo_normaldistribution: ; nidm_errorVarianceHomogeneous: "false"^^xsd:boolean ; nidm_varianceMapWiseDependence: nidm_IndependentParameter: ; **nidm_hasErrorDependence: obo_unstructuredcovariancestructure: ;** **nidm_dependenceMapWiseDependence: nidm_ConstantParameter: .**
Spatial definition	Attribute **'dependence Map-Wise Dependence'** of an **'Error Model'** entity. Possible values include: • 'Constant Parameter' for a global estimate. • 'Independent Parameter' for a local estimate. • 'Regularized Parameter' for a spatially regularized estimate.	
Contrast construction	Attribute **prov:value** of a **obo:'contrast weight matrix'** entity provides the contrast vector.	EXAMPLE: Contrast Weights @prefix nidm_statisticType: <http://purl.org/nidash/nidm#NIDM_0000123> . @prefix nidm_contrastName: <http://purl.org/nidash/nidm#NIDM_0000085> . @prefix obo_contrastweightmatrix: <http://purl.obolibrary.org/obo/STATO_0000323> . @prefix obo_tstatistic: <http://purl.obolibrary.org/obo/STATO_0000176> . niiri:contrast_id a prov:Entity , obo_contrastweightmatrix: ; rdfs:label "Contrast: Listening > Rest" ; **prov:value "[1, 0, 0]"^^xsd:string ;** nidm_statisticType: obo_tstatistic: ; # obo:'t-statistic' nidm_contrastName: "listening > rest"^^xsd:string .
Group modeling info		
Estimation method	(same as Intra-subject fMRI)	
Statistical inference Inference on statistic image (thresholding)		
Search region for analysis Location of the search space image	Attribute **prov:atLocation** of a **'Search Space Mask Map'** entity.	EXAMPLE: Search Space Mask Map @prefix nidm_SearchSpaceMaskMap: <http://purl.org/nidash/nidm#NIDM_0000068> . @prefix nidm_inCoordinateSpace: <http://purl.org/nidash/nidm#NIDM_0000104> . @prefix nidm_expectedNumberOfVoxelsPerCluster: <http://purl.org/nidash/nidm#NIDM_0000143> . @prefix nidm_expectedNumberOfClusters: <http://purl.org/nidash/nidm#NIDM_0000141> . @prefix nidm_searchVolumeInVoxels: <http://purl.org/nidash/nidm#NIDM_0000121> . @prefix nidm_searchVolumeInUnits: <http://purl.org/nidash/nidm#NIDM_0000136> . @prefix nidm_reselSizeInVoxels: <http://purl.org/nidash/nidm#NIDM_0000148> . @prefix nidm_searchVolumeInResels: <http://purl.org/nidash/nidm#NIDM_0000149> .

Volume of the search region in voxels.
 	
Volume of the search region CC.
	Attribute **'search Volume In Voxels'** of a **'Search Space Mask Map'** entity.
	
	Attribute **'search Volume In Units'** of a **'Search Space Mask Map'** entity.
		`@prefix nidm_noiseFWHMInVoxels: <http://purl.org/nidash/nidm#NIDM_0000159> .`
		`@prefix nidm_noiseFWHMInUnits: <http://purl.org/nidash/nidm#NIDM_0000157> .`
		`@prefix nidm_randomFieldStationarity: <http://purl.org/nidash/nidm#NIDM_0000120> .`
		
		`niiri:search_space_mask_id a prov:Entity , nidm_SearchSpaceMaskMap: ;`
		`rdfs:label "Search Space Mask Map" ;`
		`prov:atLocation "SearchSpaceMask.nii.gz"^^xsd:anyURI ;`
		`nfo:fileName "SearchSpaceMask.nii.gz"^^xsd:string ;`
		`dct:format "image/nifti"^^xsd:string ;`
		`nidm_inCoordinateSpace: niiri:coordinate_space_id_2 ;`
		`nidm_expectedNumberOfVoxelsPerCluster: "0.553331387916112"^^xsd:float ;`
		`nidm_expectedNumberOfClusters: "0.0889172687960151"^^xsd:float ;`
		`nidm_searchVolumeInVoxels: "65593"^^xsd:int ;`
		`nidm_searchVolumeInUnits: "1771011"^^xsd:float ;`
		`nidm_reselSizeInVoxels: "22.9229643140043"^^xsd:float ;`
		`nidm_searchVolumeInResels: "2552.68032521656"^^xsd:float ;`
		`nidm_noiseFWHMInVoxels: "[ 2.958, 2.966, 2.611 ]"^^xsd:string ;`
		`nidm_noiseFWHMInUnits: "[ 8.876, 8.898, 7.835 ]"^^xsd:string ;`
		`nidm_randomFieldStationarity: "true"^^xsd:boolean ;`
		`crypto:sha512 "e43b6e01b0463fe7d40782137867a"^^xsd:string ;`
		`prov:wasGeneratedBy niiri:inference_id .`

<u>Correction for multiple comparisons</u>
 	
Corrected or not?
Method used for correction
 	
Region over which correction for multiple comparisons was performed
	Attribute **prov:type** of the **'Height Threshold'** and the **'Extent Threshold'** used by an **'Inference'** activity
	
	Possible values include:
	• obo: 'FWER adjusted p-value' for an FWE-corrected threshold
	• 'P-Value Uncorrected' for an uncorrected threshold
	• obo:'q-value' for an FDR-corrected threshold
	
	Attribute **prov:atLocation** of a **'Search Space Mask Map'** entity.
		`EXAMPLE: Voxel-wise p<0.05 FWER-corrected threshold`
		`@prefix nidm_HeightThreshold: <http://purl.org/nidash/nidm#NIDM_0000034> .`
		`@prefix nidm_equivalentThreshold: <http://purl.org/nidash/nidm#NIDM_0000161> .`
		`@prefix obo_FWERadjustedpvalue: <http://purl.obolibrary.org/obo/OBI_0001265> .`
		
		`niiri:inference_id prov:used niiri:height_threshold_fwer_id .`
		
		`niiri:height_threshold_fwer_id a prov:Entity, nidm_HeightThreshold:,`
		`obo_FWERadjustedpvalue: ;`
		`rdfs:label "Height Threshold: p<0.05 (FWER-corrected)" ;`
		`prov:value "0.05"^^xsd:float ;`
		`nidm_equivalentThreshold: niiri:height_threshold_stat_id .`

<u>Voxel-wise significance</u> Corrected for Family-wise error (FWE) or false discovery rate (FDR)? If FWE found by random field theory list the smoothness in mm FWHM RESEL count	Attribute **prov:type** of the **'Height Threshold'** used by an **'Inference'** activity (cf. above for possible values), Attribute **'noise FWHM In Units'** of a **'Search Space Mask Map'** entity. Attribute **'search Volume In Resels'** of a **'Search Space Mask Map'** entity.	(cf. example for 'Search region for analysis' and 'Correction for multiple comparisons')

<u>Cluster-wise significance</u> cluster-defining threshold cluster significance level smoothness (for random field theory) RESEL count	Attribute **prov:value** of the '**Height Threshold'** used by an **'Inference'** activity. Attribute **prov:value** of the **'Extent Threshold'** used by an **'Inference'** activity. Attribute **'noise FWHM In Units'** of a **'Search Space Mask Map'** entity. Attribute **'search Volume In Resels'** of a **'Search Space Mask Map'** entity.	EXAMPLE: Cluster-wise p<0.05 FWER-corrected threshold with cluster-forming threshold of p<0.001 uncorrected @prefix nidm_HeightThreshold: <http://purl.org/nidash/nidm#NIDM_0000034> . @prefix nidm_PValueUncorrected: <http://purl.org/nidash/nidm#NIDM_0000160> . @prefix nidm_ExtentThreshold: <http://purl.org/nidash/nidm#NIDM_0000026> . @prefix obo_qvalue: <http://purl.obolibrary.org/obo/OBI_0001442> . niiri:extent_threshold_fdr_id a prov:Entity, nidm_ExtentThreshold:, obo_qvalue: ; rdfs:label "Extent Threshold: p<0.05 (FDR-corrected)" ; prov:value "0.05"^^xsd:float . niiri:height_threshold_unc_id a prov:Entity, nidm_HeightThreshold:, nidm_PValueUncorrected: ; rdfs:label "Height Threshold: p<0.001 (uncorrected)" ; prov:value "0.001"^^xsd:float .

Table 2. Checklist to report neuroimaging results for intra-subject fMRI and group models from (Poldrack et al. 2008) and corresponding representation in NIDM-Results. The following items from the original checklist were excluded as not available automatically: "design type", "orthogonalization of regressors", "additional regressors used","if not whole brain, state how region of analysis was determined", "If correction is limited to a small volume, the method for selecting the region should be stated explicitly", "threshold used for visualization in figures", "correction for multiple planned comparisons within each voxel".