NeuroImage 82 (2013) 662–670

Contents lists available at ScienceDirect

NeuroImage

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

Review

Modeling functional Magnetic Resonance Imaging (fMRI) experimental variables in the Ontology of Experimental Variables and Values (OoEVV)

Gully A.P.C. Burns [a,*], Jessica A. Turner [b,c,d]

[a] Information Sciences Institute, University of Southern California, 4676 Admiralty Way, Suite 1001, Marina del Rey, CA 90292-6695, USA
[b] Mind Research Network, 1101 Yale Blvd NE, Albuquerque, NM 87106, USA
[c] Dept of Psychology, University of New Mexico, Albuquerque, NM 87106, USA
[d] Dept of Psychiatry, University of New Mexico, Albuquerque, NM 87106, USA

ARTICLE INFO

Article history:
Accepted 3 May 2013
Available online 16 May 2013

ABSTRACT

Neuroimaging data is raw material for cognitive neuroscience experiments, leading to scientific knowledge about human neurological and psychological disease, language, perception, attention and ultimately, cognition. The structure of the variables used in the experimental design defines the structure of the data gathered in the experiments; this in turn structures the interpretative assertions that may be presented as experimental conclusions. Representing these assertions and the experimental data which support them in a computable way means that they could be used in logical reasoning environments, i.e. for automated meta-analyses, or linking hypotheses and results across different levels of neuroscientific experiments. Therefore, a crucial first step in being able to represent neuroimaging results in a clear, computable way is to develop representations for the scientific variables involved in neuroimaging experiments. These representations should be expressive, computable, valid, extensible, and easy-to-use. They should also leverage existing semantic standards to interoperate easily with other systems. We present an ontology design pattern called the Ontology of Experimental Variables and Values (OoEVV). This is designed to provide a lightweight framework to capture mathematical properties of data, with appropriate 'hooks' to permit linkage to other ontology-driven projects (such as the Ontology of Biomedical Investigations, OBI). We instantiate the OoEVV system with a small number of functional Magnetic Resonance Imaging datasets, to demonstrate the system's ability to describe the variables of a neuroimaging experiment. OoEVV is designed to be compatible with the XCEDE neuroimaging data standard for data collection terminology, and with the Cognitive Paradigm Ontology (CogPO) for specific reasoning elements of neuroimaging experimental designs.

Contents

Introduction

As a discipline, neuroimaging has several natural characteristics which promote the ease and impact that informatics and data-sharing have for the field. Imaging systems automatically generate large quantities of digital information as part of their normal function. Industry

* Corresponding author.
E-mail addresses: gully@usc.edu (G.A.P.C. Burns), jturner@mrn.org (J.A. Turner).

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.05.024

standards such as the DICOM image format provide access to imaging meta-data appropriate for the medical requirements of the imaging methods. Software and data sharing architectures (such as NITRC, www.nitrc.org) provide access to community-developed analysis methods, data repositories, text- and data-mining and ontologies. These provide a powerful set of resources that enable users to share, analyze and consolidate data across many different experiments.

A key issue to address in how these experimental data contribute to understanding brain function is capturing the descriptive meta-data that determine the interpretation of experiments. An investigator attempting to determine the functional Magnetic Resonance Imaging (fMRI) evidence for dysfunctional auditory processing in schizophrenia, for example, would know implicitly which types of experimental paradigms are relevant to the study of schizophrenia and what types of measurement indicate the presence or absence of an effect in data. Informatics systems that are either only focussed on sharing the underlying data itself or the high-level interpretations of what the data mean do not explicitly represent this crucial knowledge. This paper is concerned with developing an informatics-driven framework that can provide a foundation for connecting experimental results (e.g., similar statistically-significant differences in blood oxygenation level dependent, BOLD, signal across different experiments) with their interpretation (e.g., the dysfunction in auditory cortex in subjects with schizophrenia).

The precise definitions of the *variables* forming the basis of experimental designs are central to this process of linking data to studies' conclusions. These variables include standard and study-specific clinical assessments, study-specific experimental conditions, patient demographics, simple control variables, and exclusionary conditions (such as history of drug or alcohol abuse) as well as the underlying measurements made in the study (such as cortical BOLD signal values). The investigation of brain function implicitly involves combining these variables into experimental designs that demonstrate and examine observable statistical effects. The choice of variables (and their analysis) determines the relevance and effectiveness of the experiment to answer underlying questions about brain function.

These variables are often not included in standard formats such as DICOM, which focuses on the subject-specific imaging protocol and very basic information regarding the image, or in data annotation terminologies such as XCEDE (Gadde et al., 2012) which focus on describing the data set per se, its characteristics and provenance, but not the larger experimental context in which the data were collected. XCEDE and ontologies such as the Cognitive Paradigm Ontology, CogPO (Turner and Laird, 2012), provide concrete frameworks for a subset of the needed definitions, but do not explicitly provide an underlying ontological framework that can support other, previously unseen variables. For example, an fMRI study of subjects with schizophrenia doing an auditory oddball task generates DICOM images from the MRI scanner. The DICOM data include meta-data about the imaging parameters and subject orientation, but not information about the timing of the different auditory stimuli, or about whether the participant had schizophrenia. The XCEDE meta-data could include the task timing and behavioral information, as well as the subject diagnosis, and even how the analyses of the fMRI data were done. This greatly improves data sharing; but the XCEDE representation is not designed to link data and their interpretations about dysfunctional auditory processing in schizophrenia. The published paper provides that link, but in a way that is currently only interpretable and re-usable by humans. Making that link explicit is only possible if we create a practical, extensible and flexible representation of the experimental variables used in imaging experiments that may then be explicitly linked to interpretive statements about brain function.

At one end of the informatics continuum is idiosyncratic data tagging (e.g. “AO” for one lab means that the data were collected with a specific auditory oddball task, while someone else uses it to mean “age of onset” for a disease). XCEDE represents a step beyond that minimum (e.g. providing common definitions for the tags, repositories of scan type definitions, and graphical recreations of data pre-processing steps). NeuroLex provides a step beyond XCEDE in providing a broader, loosely hierarchical representation of scan types and their subtypes, for example. Then, at the other end of this continuum is the full ontological representation, which carefully draws distinctions between the concept of an MRI and the particular MRI scanner used in an experiment, or between the plan to perform an experiment and the actual experiment itself (see Brinkman et al., 2010). This level of ontological rigor is necessary eventually, to avoid logical errors in broader reasoning; but it is not immediately necessary for the purposes of improving the current semantic frameworks for representing neuroimaging experiments. For the moment, we attempt an incremental approach, with a basic semantic framework for experiments that can be linked into the more formal ontologies and expand those as this grows.

By making variables the focus of our representation, we seek to bridge the gap between simple data mark-up for sharing and re-using, and the complexities of a full ontology-driven approach to neuroimaging. This is the underlying strategy of our approach called ‘Knowledge Engineering from Experimental Design’ (KEfED, Russ et al., 2011; Tallis et al., 2012). Within KEfED, we focus on the dependency of dependent variables (measurements) and independent variables (parameters and constants) within the design of a single experiment, so that we can construct representations of the experiments and results that reflect that dependency. The way that we construct neuroimaging experiments to reveal informative patterns in data is based on these subtle dependencies and comparisons between cases that have, until now, not been explicitly represented in any informatics system.

As a starting point for developing flexible, accurate repositories for neuroimaging knowledge, we define a practical system for specifying experimental variables used for neuroimaging experiments. This central role of experimental variables in our KEfED methodology prompted us to develop an ‘ontology design pattern’ (ODP) for the semantic elements that make up the core representation used in our approach. Analogous to design patterns in software engineering, ODPs are reusable components that may be used to assist with the construction of full, more detailed ontologies (Blomqvist, 2009; Gangemi and Presutti, 2009). This approach has being applied to the development of ontologies in biomedicine (Aranguren et al., 2008) and provides a way of developing small-scale, pragmatic, reusable, components. We specifically intend to use the ODP approach to provide lightweight ontological definitions of variables and their values that be easily used and reused by research scientists (including neuroimaging specialists, Burns and Turner, 2012).

Note that at this early stage, this approach seeks only to represent ‘just enough’ detail to capture the dependency relationships that exist between independent and dependent variables for use within the KEfED approach (Russ et al., 2011). We differentiate our efforts from a formal, top-down ontology development process by presenting our work as an ODP. Our system is intended to provide a preliminary way of defining only the bare minimum of ontological detail that may then be further developed by more comprehensive ontological approaches if necessary. This framework is intended as being interactive with other ontological efforts (such as CogPO, XCEDE, or the Ontology of Biomedical Investigation).

In this paper, we introduce this methodology by applying it to neuroimaging variables in a specific example study. Our intention here is to direct neuroimaging researchers' attention to the definitions of experimental variables as a viable starting point for developing consolidated knowledge repositories that are capable of seamless data sharing. The complexities of carefully-constructed, logic-driven representations crafted by ontologists are often difficult to use by neuroimaging researchers directly. We provide a more manageable starting point: definitions of the parameters and measurements used in neuroimaging experiments. This process of developing ontologies for individual research domains is a capability that has been highlighted in the core work of the Knowledge Engineering Working Group of the Biomedical Informatics Research Network (BIRN, Helmer et al., 2011).

Methods

Conceptual design

An abbreviated version of the high-level design of the OoEVV model (0.0.8) is shown in Fig. 1. This shows the core *OoevvElementSet* class that defines groupings of variables, entities and processes to form the raw material for a KEfED model within a specific domain (see Russ et al., 2011). Each example of an *OoevvElementSet* can contain several *OoevvElements* (which must be one of the three possible subtypes: *ExperimentalVariable*, *OoevvProcess*, or *OoevvEntity*). Within this paper we are primarily focussed on the definition of relevant experimental *variables* pertaining to neuroimaging and so we will not elaborate further on *entities* (e.g., a brain, a subject or a scan) or *processes* (e.g., collecting the fMRI scan, running a GLM analysis). For the purpose of connecting to the larger ontological community, the *OoevvProcess* class may be mapped onto the 'planned process' class from the Ontology of Biomedical Investigation ('OBI', obo:OBI_0000011) and the *OoevvEntity* class onto either the 'material entity' class of the Basic Formal Ontology ('BFO', obo:BFO_0000040) or the 'information content entity' from the Information Artifact Ontology ('IAO', obo:IAO_0000030).

Each *ExperimentalVariable* instance must have a single specified *MeasurementScale*. Data values (e.g., the subject's age or diagnosis, the median activation in BA 41) are denoted as *MeasurementValue* instances which each must be associated with a *MeasurementScale* instance.

The *OoevvElementSet* construct allows us to group related variables together without formally defining the criteria under which they may be grouped. This grouping could be based on a study, a research program made up of multiple related studies, or an entire biomedical field such as 'fMRI studies in general'. This structure is provided to permit the easy reuse of the ODP in the context of a specific scientific application. The cardinality constraints on the association with the *ExperimentalVariable* class indicate an 'n-to-n' relationship, meaning that (a) each set contains several elements and (b) each element may be included in several sets.

Each variable 'measures' a 'Quality' defined in an external ontology. Most notably, this could refer to terms from the 'Phenotype Ontology', PATO (Mungall et al., 2010), but could also refer to physical qualities, such as 'temperature' (obo:PATO_0000146), or derived measures such as the values in a contrast image from a GLM analysis of the subject's fMRI data. The *Quality* element shown in Fig. 1 provides a structure to capture the semantic meaning of the variable and inform interpretations based on measurement values. The presence of this link permits us to define multiple variables that measure the same underlying quality (such as temperature measurements based on the Fahrenheit or Celsius scales).

The scale construct is where we have placed most of our design effort, intending that specializations of the *MeasurementScale* class provide the computational structure for processing of different types of measurement values (see Table 1). In the same way that a computer language must distinguish between integers, floats and strings, our system must differentiate between different types of measurement relevant to neuroimaging experiments. Each scale type provides a design specification, providing the information needed to determine whether a given value would be a permitted value for a specific scale. We illustrate this with neuroimaging examples in the Results section.

The underlying rationale for these distinctions is that measurement scales provide a framework for the mathematical and computational properties of data that is largely divorced from the meaning of different data points. The original formulation on a theory for measurement scales named four such basic types: nominal, ordinal, interval and ratio (Stevens, 1946). This allows a distinction between a grouping or nominal variable, for example, such as "Patient" or "Control", and a stratification or ordinal variable such as "Severely Impaired", "Mildly Impaired", and "Not Impaired." These non-quantitative scales are distinguished from numerical scales such as "Age in Years", where computations such as taking a weighted average or a ratio can be meaningful.

In our current model, our focus is more based on practical computational concerns than strict mathematical definitions and so do not distinguish between interval and ratio scales (instead we use the simple designation 'numeric' to specify a quantitative measurement scale, with three particular subtypes to capture the different computational properties of the values; see Table 1). For measurements that have more complex data structure, we provide an application-programming interface (API) for OoEVV that would permit developers to extend these core types by defining their own scales within the underlying software architecture of the OoEVV system.

Neuroimaging applications frequently use 2-, 3- and 4-dimensional data sets that are usually embedded into data files of a specified format (such as JPG, DICOM, and NIFTI). We define a measurement scale subtype (*FileScale*) that specifies its values simply as files of a given format. This allows us to define variables (such as 'Original fMRI data set') whose values are given by the data files themselves.

There is a close correspondence between these base elements of OoEVV with high-level terms from the Ontology of Biomedical Investigation (OBI, Brinkman et al., 2010). In broad strokes, the *ExperimentalVariable* class in OoEVV most closely corresponds to the 'directive information entity' term (IAO_0000033); the *MeasurementScale* class corresponds to the 'value specification' term (definition currently under discussion); the *MeasurementValue* class corresponds to the 'measurement datum' term (obo:IAO_0000109). We directly refer to

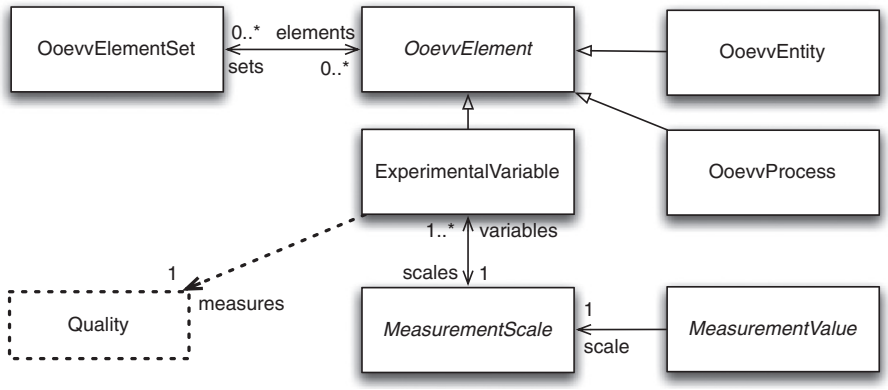

Fig. 1. High-level UML design of core classes in OoEVV. Solid rectangles are UML class definitions within OoEVV. Arrows show navigable associations between classes with roles and cardinality values. Dotted rectangles and arrows designate class definitions and associations linked to external ontological resources.

Table 1
Different types of scale modeled for use within OoEVV.

Scale type	Characteristics
BinaryScale	True/false
BinaryScaleWithNamedValues	A BinaryScale where the true and false values are named (e.g., present/absent)
NumericScale (abstract)	Numeric scales may have preset maxima and minima, they may be 'circular' (e.g. angles in radians) and may have defined units.
DecimalScale	An extension of NumericScale with real values.
IntegerScale	An extension of NumericScale with integer values.
TimestampScale	An extension of NumericScale with timestamped values.
HierarchicalScale	Hierarchically organized values (e.g., a value from a phylogenetic taxonomy)
NominalScale	Values are names (e.g., 'patient', 'healthy control'); also known a categorical scale
NominalScaleWithAllowedTerms	Values are names that must be drawn from a predefined list
OrdinalScale	Values are ranks (e.g., symptom severity scores)
OrdinalScaleWithMaxRank	Values are ranks with a predefined maximum
OrdinalScaleWithNamedRanks	Values are ranked in order where each value is a term (e.g., infant, adolescent, adult, senescent)
CompositeScale	This scale is made up of several sub-variables. (e.g., a neuropsychological assessment may contain several questions, each one corresponding to a well-defined variable with an ordinal scale.)
RelativeScale	Values are defined by relations to a set of predefined terms (i.e., a volume of brain tissue may be defined by overlap/inclusion relations with named brain structures, Russ et al., 2011).
FileScale	This scale denotes that the measurement is itself contained within a computer file.
NaturalLanguageScale	Values are passages of Natural Language text in a specified language (e.g. free text responses to interview questions)

the 'quality' term from the Basic Formal Ontology (obo:BFO_0000019) to specify the quality element referred to in Fig. 1 (and will use a local representation that we submit to PATO for curation when a suitable quality cannot be found). The *OoevvElementSet* class has no equivalent in OBI.

Software implementation

We use the 'View Primitive Data Model framework' (VPDMf) to provide the architecture for the underlying logical design of OoEVV and for the curation system involved in populating and using the model. The VPDMf provides a programming interface permitting operations on 'views' defined over a schema defined in Unified Modeling Language (UML). Each view is expressed as a graph over the classes of the model so that we may perform data management operations on 'view instance' entities which contain connected class instances as encapsulated graph structures.

As a base representation, the UML provides a framework that may be considered a subset of those available under OWL/RDF. These include class instances, class inheritance, class attributes and associations with domain, range and cardinality constraints. In addition, the VPDMf provides a powerful scaffolding framework that permits the underlying model classes to be directly generated in Java, Actionscript, and OWL with corresponding tables in MySQL (with some programming elements added such as primary and foreign keys and set backing tables). Industrial-strength programming tools support software engineering within applications built on the VPDMf (including in-line highlighting of compilation errors and automated error resolution in the Eclipse development environment, powerful library dependency tools such as Maven). This provides a basis for agile development of specialized tools that facilitate OoEVV curation.

Data curation process

A goal of OoEVV is to provide a curation framework that neuroimaging experts could use easily. We developed a command-line application that uses Excel Workbooks to curate terminology. The data flow between the various components of this system is shown in Fig. 2.

As shown, the system provides three methods for users to generate OoEVV-formatted artifacts for use within the system (empty spreadsheets, an un-instantiated OWL file or an empty MySQL database). The starting point for most curation is the *generateOoevvSpreadsheet* command to create a formatted 'OoEVV Spreadsheet' as an Excel workbook to be populated. Each separate OoEVV spreadsheet file corresponds to a separate *OoevvElementSet*. The freedom provided by using simple spreadsheets inevitably leads to errors (through non-conformance to the curation procedures and edge-cases that break the system). Processing these spreadsheets can be performed through commands to (a) upload a single spreadsheet to an OWL file (*ooevvSpreadsheetToOwl*); to (b) upload a single spreadsheet to the database (*ooevvSpreadsheetToDatabase*); or to (c) upload all the spreadsheets found in a directory to the database (*ooevvDirToDatabase*). The system also includes an aggregation mechanism (the *databaseToOwl* command) where the system iterates over the entire OoEVV representation in the database to generate an OWL model derived from all spreadsheets. This OWL file may then be uploaded to the National Center of Biomedical Ontology's (NCBO) BioPortal system to provide a centralized, versioned representation of OoEVV for global use (http://bioportal.bioontology.org/ontologies/3006).

Results

We focused as an example experiment on the paper by Ford et al. (2009) "Tuning into the Voices: A Multisite fMRI Study of Auditory Hallucinations." In this study, subjects with schizophrenia and age- and gender-matched controls were recruited as part of the FBIRN Phase II study across several universities. Their years of education and two measures of socioeconomic status (Hollingshead and Redlich, 1958), and handedness (left, right, or ambidextrous) as determined by the Edinburgh Handedness Inventory were explicitly reported within the manuscript. Subjects were scanned while performing an 'auditory oddball task' (from the CogPO ontology; the term ID is COGPO: COGPO_00072) consisting of streams of tones of a standard frequency, with the rare 'oddball' tone of a different frequency which subjects were supposed to respond to with a finger press. The primary dependent variable was the median contrast between BOLD signal response to the oddball or target tone relative to the standard tone in several auditory and visual cortex regions of interest (ROIs). The first analysis was an Analysis of Variance (ANOVA) on this median contrast measure in healthy controls (HC) vs. subjects with schizophrenia (SZ), including effects of diagnosis, hemisphere, and ROI, and interactions among these factors. The second analysis was a similar ANOVA but compared subjects with schizophrenia with a report of recent hallucinations (based on their answers to a particular question on the 'Scale for the Assessment of Positive Symptoms', or SAPS, Andreasen, 1984) to subjects with schizophrenia and no hallucinations. A third ANOVA was the same as the second but included behavioral accuracy on the oddball task as a covariate.

The most succinct statement of the results from these analyses is in the abstract, which states, "Hallucinators had less activation to probe tones in left primary auditory cortex (BA 41) than nonhallucinators." The Results section identifies the main effects and interactions of each of the factors in the first ANOVA, identifying that HC subjects showed greater median activation values than SZ subjects overall, different

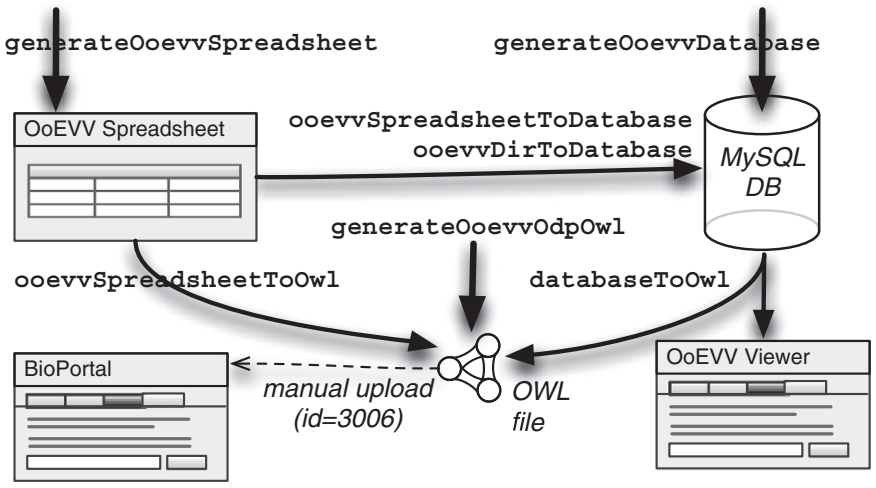

Fig. 2. Dataflow provided by the OoEVV biocuration process. (See text for more detail.)

ROIs showed greater or lesser effects, the left hemisphere was more active than the right, and there was a group × ROI interaction in that the difference between BA 22 and BA 21 was greater in HC than SZ. The second ANOVA is reported similarly, including a 3 way group × hemisphere × ROI interaction—nonhallucinating SZ subjects had greater activation in BA 41 on the left than did hallucinating SZ, and this was not true in the right hemisphere. This effect was not changed by using accuracy as a covariate in the 3rd ANOVA, and is the focus of the Discussion section. The final assertions from the paper are, in effect "Subjects with schizophrenia activate less to the oddball tone than do healthy subjects, in the auditory cortex" and "Subjects with schizophrenia who hallucinate activate less to the oddball tone in BA 41 than do those who do not hallucinate."

Our eventual goal within the KEfED approach is to provide a general-purpose computational representation for these statements. In fact, existing work (reported in Russ et al., 2011) provides the base representation for a given for a data point or mean value based on the set of parameters associated with that particular measurement. Since the statements described above involve a comparative statement between two data points (as a statistical relation between them), capturing them would require an extension of the KEfED model, which is beyond the direct scope of this paper. Our focus here is on the raw material for such a representation: a representation of the constituent variables and their values needed to express these findings.

To capture the variables for this experiment and analyses in OoEVV, we focused on variables reported explicitly in the paper or used in the analyses. These include subject variables such as age, diagnosis, symptom severity scores, and the subject's accuracy on the task; data collection variables such as the fMRI methods and experimental conditions; and data analysis variables such as the pre-processing steps, and the regions of interest. All factors and covariates included in the final ANOVAs must be represented in OoEVV; however, for the purposes of representing the experimental context, variables that are not explicitly included in the analysis (i.e., age, gender, handedness, SES scores, other clinical assessments, the experimental conditions, scanner make and model) should also be able to be presented within the OoEVV framework. The framework for representing statistical analyses, to link the final interpretation to the relevant F-statistic, is not yet included; however, the variables and factors and the structure of the experiment leading into the ANOVAs can be represented.

In the tables below, we present examples of the variables from the experiment, with links to the relevant qualities they measure, the definitions of those qualities, and the scales they are drawn from. Every variable has a scale, and a quality, which it measures. The scale for each variable is an OOEVV term, and is a subclass of one of the scales listed in Table 1. Examples of allowable values for specific scales are included in the final table.

Additionally, the tables list the ontology ID and preferred name of the pre-existing ontology term that denotes the quality that the variable is measuring (when such a definition is available from the NCBO BioPortal site: bioportal.bioontology.org). The NCBO BioPortal lists many biomedical ontologies and allows cross-ontology searching and visualization to find the best match for the intended concept. Several of the variables, such as the BOLD signal contrast values from the subject-level analysis of the oddball fMRI data, do not currently have definitions in any existing ontology. By providing OoEVV as an initial framework to capture, define and publish these elements in the lightweight OoEVV representation in BioPortal, we can provide a representation that neuroimagers may both use and contribute to straightforwardly.

As can be seen in Tables 2 and 3, the qualities being measured have largely been used and defined in at least one other ontology. In OoEVV, we pull as much as possible from ontologies which are already in production, though this leads to a challenge in identifying which ontology should be selected for a particular term. For example, the diagnosis of either 'schizophrenic' or 'control' used in the Ford sample was made on the basis of a Structured Clinical Interview for Diagnosis (SCID, Michael et al., 2002) based on the DSM-IV-TR categories. In annotating the FBIRN dataset itself, the variable based on the SCID questions should be linked to some standard term for the quality being measured (which could be thought of as the quality for an individual of having being diagnosed as schizophrenic). The NIFSTD ontology includes terms for SCID, and 'schizophrenia', but not for 'diagnosis'. The Ontology of General Medical Science (OGMS) includes 'diagnosis' but not the SCID, or DSM-IV. Searching NCBO's BioPortal for 'diagnosis' returns 164 results from 33 separate ontologies. Choosing which term to link to is a function of several criteria: the ontology's level of specificity (cancer or pediatric disease ontologies are too specific, while the MeSH is too general), its completeness (does the term include a definition, or links to synonyms in other ontologies?), and its correctness for the intended use within OoEVV.

The quality a variable is referring to is separate from the allowable values on the scale that measure it. The term 'diagnosis' could be measured using a list of ICD-10 codes, the DSM-IV-TR or the eventual DSM-V categories (each one requiring a different scale), or other standard terminologies of disease and disorders. By explicitly representing the DSM-IV-TR scales as a defined entity, we capture the Ford et al. researchers' decision to use the DSM-IV-TR codes in their work directly.

The 'Scanner-type' scale uses a *NominalScale* (meaning that the scale simply reflects a set of named values so that two values may be considered equal only if the names are the same), while the 'behavioral-paradigm-name' scale uses a *NominalScaleWithAllowableValues*

Table 2
Example subject variables from the Ford et al. experiments. Abbreviations used in the table for full names of ontologies are as follows: BRO: Biomedical Resource Ontology; CogPO: Cognitive Paradigm Ontology; EFO: Experimental Factor Ontology; FMA: Foundational Model of Anatomy; NCIt: NCI Thesaurus; NIFSTD: Neuroscience Information Framework Standard ontology; OGMS: Ontology of General Medical Science; PATO: Phenotype, Attribute, and Trait Ontology; SNOMED: Systematized Nomenclature of Medicine.

Variable name	Definition	Ontology	Measures	Scale name
Subject's-age	The age of the subject participating in the experiment (in years)	PATO	PATO: age A time quality inhering in a bearer by virtue of how long the bearer has existed.	Time-duration
DSM-IV measurement	Diagnosis as determined by Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, (DSM-IV)	OGMS	OGMS: diagnosis The representation of a conclusion of a diagnostic process	DSM4-scale
Edinburgh Handedness Inventory	A value from −100 to +100 denoting the handedness of the subject (the inventory involves a questionnaire and computes a score based on that value).	NIFSTD	NIFSTD: handedness A behavioral quality inhering in a bearer by virtue of the bearer's unequal distribution of fine motor skill between its left and right hands or feet.	Edinburgh-handedness-inventory-scale
Socioeconomic Status Questionnaire	Socioeconomic Status as measured by the Hollingshead and Redlich (1958) Questionnaire	NCIt	NCIt: socioeconomic factors Characteristics of a person such as education and occupation, used to describe the person's position in stratification systems, and access to services	SES-scale
Global rating of severity of hallucinations	The global severity of hallucinations from the SAPS	SNOMED	SNOMED: hallucinations Perceptions in a conscious and awake state in the absence of external stimuli which have qualities of real perception, in that they are vivid, substantial, and located in external objective space.	6-Point-severity-with-unknown-option
Performance accuracy	The behavioral accuracy of the subject on a task	NCIt	NCIt: accuracy The quality of nearness to the truth or the true value.	Percent-correct-scale
Site of testing	The physical location the data collection were performed at	EFO	EFO: site A site is an entity which consists of a characteristic spatial shape in relation to some arrangement of other material entities.	Site-name
MRI scanner	Brand and type of MRI scanner used for data collection	BRO	BRO: MRI_Scanner An image acquisition device that visualizes detailed internal structure and limited function of the body, using a powerful magnetic field to align the nuclear magnetization of (usually) hydrogen atoms in water in the body.	Scanner-type
fMRI behavioral experiment	The cognitive task used during the fMRI scans	CogPO	CogPO: behavioral experimental paradigm The behavioral aspects of the experiment: what stimuli are presented to the subject when, and under what conditions, and what the subject's responses are supposed to be.	Behavioral-paradigm-name
Original fMRI data set	The fMRI data from a single subject's scanning session	NIFSTD	NIFSTD: Blood Oxygenation Level Dependent signal	Image-file-scale
Pre-processed fMRI data set	The output of pre-processing the original fMRI data, to do motion correction, slice timing correction, etc.	OOEVV	OOEVV: pre-processed BOLD signal The BOLD signal following motion correction, normalization, etc. prior to analysis.	Image-file-scale
fMRI contrast image	The spatial image containing the values contrasting various conditions from a regression analysis on processed fMRI data	OOEVV	OOEVV: contrast-image values The contrast of beta-values from an fMRI GLM analysis	Image-file-scale
fMRI median activation	The median value of a contrast image within a particular mask or region	OOEVV	OOEVV: median activation value The median value of cortical activation	Arbitrary-numericScale
fMRI ANOVA	The output of a statistical analysis comparing fMRI-activation under different conditions	OOEVV	OOEVV: significance-of-ANOVA The significance of the change in BOLD signal	ANOVA-results
Brodmann area label	The identification of a specific brain region in a sample as corresponding to a region drawn from Brodmann's parcellation scheme	FMA	FMA: Brodmann_Area (no definition provided)	Cortical-region-Brodmann-scale

(indicating that the possible names must be drawn from a restricted set). This situation reflects how behavioral paradigm names are drawn from a predefined list in CogPO but the make and model of scanners are an open-ended set, with new scanner names being created every year; thus OoEVV does not pull from a pre-defined list of values for that scale. The curator may choose different levels of stringency for each variable depending on the requirements of their application and of their knowledge of the domain.

The measurement scales of variables with a well-defined substructure are represented by the *CompositeScale* construct. This is a measurement scale that contains other, simpler variables (through a 'partsOf/hasParts' property). One example of this is the ANOVA-results scale, that contains sub-variables to store the statistical results of the ANOVA analysis (such as the F statistic and p-values). *CompositeScale* instances are also used to represent the substructure of questionnaires (such as the Abnormal Involuntary Movement Scale or 'AIMS', http://

Table 3
Example scales and allowable values from the fMRI variables shown in Table 2.

Scale name	Full name	Definition	Scale type	Units	Allowable values
Time-duration	Time duration	The duration of a time interval	DecimalScale	Years	
DSM4-scale	DSM4 code	A nominal scale generated from the DSM4 structured interview	NominalScaleWithAllowedValues	none	The list of DSM-4 diagnoses and subdiagnoses
Edinburgh-handedness-inventory-scale	Edinburgh Handedness Inventory Scale	An ordinal scale computed from the Edinburgh Scale where −100 means the person is strongly left handed, and +100 indicates they are strongly right handed	OrdinalScale	none	−100 to +100
SES-scale	Socioeconomic status scale	The score on specific questions from the Hollingshead questionnaire	OrdinalScale	None	1–7
6-Point-severity-with-unknown	Six point ordinal scale for severity with opt-out code	An ordinal scale denoting the severity of a symptom; includes a "not known" response	OrdinalScaleWithNamedRanks	None	
Site-name	Site name	The code for the site where the MRI data were collected	NominalScale	None	
Scanner-type	MRI scanner type	The scale of allowable scanner types	NominalScale	None	Siemens Tim Trio, Siemens Avanto, GE Edge, etc.
Behavioral-paradigm-name	Behavioral-paradigm-names	A nominal scale for the various types of fMRI tasks or experiments	NominalScaleWithAllowedValues	None	The subclasses of CogPO: behavioral experimental paradigm
ANOVA-results	ANOVA results table	A composite scale including the F values and significance of the effects tested in the ANOVA	CompositeScale	None	
Cortical-region-Brodmann-scale	Brodmann Region Scale	The listing of Brodmann areas	NominalScaleWithAllowedValues	None	All the Brodmann area labels

www.cqaimh.org/pdf/tool_aims.pdf) where each question could itself be considered a separate variable (*e.g.*, question 1 of the AIMS provides the definition of a variable 'the ordinal severity of involuntary movement in muscles of facial expression').

With these variables, scales, composite scales, and allowable values included in OoEVV, the process of specifying the experimental design of fMRI papers within a curation task can be facilitated. While each new experiment will have its own combination of variables, each requiring new subclasses of scales in some cases, the ability to capture subject specific variables, data collection variables, and analysis inputs and outputs provides the groundwork for a general representation. Importantly, we strive to be descriptive, rather than prescriptive and provide tools that may be used by neuroimagers to capture only the logic of how variables are linked to the underlying data. In this case, the representation uses the design pattern formalism to provide a simple framework that may express standard measurements in a uniform way.

Discussion

The central role of scientific variables in the representation described in this paper reflects the novelty (and potential impact) of our approach. As an information construct, scientific variables are ubiquitous within experimental and modeling work; they provide the basic mathematical 'hooks' for data processing and analysis but are not themselves typically the subject of detailed scrutiny within biomedical informatics systems (with some exceptions, described below). The types of variables used within an fMRI neuroimaging study include the data surrounding the fMRI signal itself, but also the variables forming the overall design of the study (such as the definition of experimental groups based on neuropsychological assessment data and the selection of controls). The selection and definition of these meta-data form the basis for the underlying interpretive *meanings* of any assertions generated within each study. Despite there being widely-accepted standard publications and practices within the field, there are no well-defined, practical ontologies that provide standardized representations of experimental observations in terms of these variables. This paper describes a preliminary attempt to formulate such an approach for neuroimagers to capture and represent observational assertions from papers and to construct practical tools to realize it.

The OoEVV project has arisen within the broader context of an approach called 'Knowledge Engineering from Experimental Design' (KEfED, Russ et al., 2011) designed to enable the creation of observational assertions based on the values of experimental variables (possibly as 'nanopublications': small-scale elements such as statements about the experiment or the interpretation of the results, that would themselves be citable pieces of knowledge with associated authorship and other provenance) (Groth et al., 2010). The KEfED formulation has, so far, demonstrated its ability to provide a generally applicable framework for direct measurements (and simple descriptive statistics such as means).

The application of the KEfED methodology in another field (the study of HIV vaccines) illustrates how curating KEfED models from different studies permit researchers to compare experimental designs between experiments with different protocols (Tallis et al., 2012). In that application, representing vaccination studies in KEfED, we were able to identify the similarities and differences in experimental methods, materials, and parameters across multiple experiments. This allows the analysis of dependencies between study variables and their outcomes, potentially identifying causes for disagreements in the literature, or experimental combinations of variables which still need to be tested. By applying this framework to neuroimaging, we provide a framework in which we can compare neuroimaging results captured with different experimental protocols, and plan future experiments based on the way that the variables' definitions relate to the underlying phenomenon of brain function. A needed addition to the framework to achieve these goals is of course the ability to represent the more complex data-driven assertions based on statistical significance between an experimental and control populations for a given measurement variable. This basic structure could provide a general framework for neuroimaging nanopublications concerning neuroimaging studies.

Within the neuroimaging community there are many projects which seek to facilitate not just the sharing of datasets, but information aggregation across studies—the coordinate-based representations of BrainMap (Fox et al., 2005; Laird et al., 2005) and Neurosynth

(Yarkoni et al., 2011) are two of the most helpful. BrainMap is a repository of results from the fMRI and PET literature; it uses a specific schema to represent information about the experimental contrasts and direction of results for every set of coordinates reported in a neuroimaging publication. It is an expert-curated results repository which focuses specifically on annotating group-level results, in contrast to data repositories which need to annotate datasets at the level of the individual subject. Many of the tables and fields in the BrainMap schema capture the experimental variables (as we understand the process in the OoEVV project)— they include the conditions of the experiment, specific characteristics of the subject sample, the context of the study, and the MRI scanner field strength, among others. Structuring the information in this way has made the BrainMap repository ideal for retrieving imaging results for meta-analysis with many successful contributions to the field (Glahn et al., 2005; Laird et al., 2009, 2011; Minzenberg et al., 2009). What it does not do is link to semantic information or structured information which can be used by automated sources, to identify when, for example, a study of 'schizophrenia' used DSM-IV categories to select its subjects or when it did not; nor could it incorporate the analyses reported in the Ford et al. paper, which are not summarized as maximal voxels from a particular contrast. However, just as the BrainMap-derived ontology of CogPO provides many standard variables for OOEVV's representation of fMRI data collection methods, the BrainMap schema provides a selection of variables which when linked to a semantic representation will give the neuroimaging community a distinct advantage in representing and reasoning across nanopublications.

The neurosynth.org website is a more recent text mining effort across neuroimaging papers, which automatically identifies tables of brain coordinates within each paper, and can tag those coordinates with important words or phrases from the text of the paper. This allows the very informative collation of which areas in the brain are likely to be tagged with various imaging terms, cognitive process labels, or experimental task names. The automatic nature of the tag extraction and annotation has the benefit of being much faster than expert curation and unbiased by human interpretation. It has the disadvantage that the tags are not always semantically related to each other—the tags allow retrieval of results, but not the next step of relating variations in experimental design or analysis to the results, for example.

The Cognitive Paradigm Ontology (CogPO, Turner and Laird, 2012) constructs a well-defined ontological representation that mirrors the structure of variables used within the BrainMap database (Fox et al., 2005). The particular choice of variables selected to characterize fMRI experiments within CogPO (stimulus, response, instruction) defines a specific set of parameters of high importance to characterizing cognitive tasks within fMRI experimental designs, therefore provides a well-defined encoding of a specialized measurement scale within the framework provided by OoEVV. A notable development is the publication of standard 'Common Data Elements' defined by the National Institute of Neurological Diseases and Stroke (NINDS, Saver et al., 2012) as well-defined, simple standard definitions for variables. The development of Minimum Information checklists (specifically, the 'Minimum Information about an fMRI Study', Poldrack et al., 2008) provides timesaving guidelines, data models and spreadsheets for researchers to follow. These work well in conjunction with standard ontologies to help shape the framework for data curation within the subject.

The XCEDE structures are being built to capture details of the pre-processing and analysis procedures (see the Keator et al. paper in this issue), which will flesh out the details of the fMRI data variable structures in OoEVV as that develops. OoEVV is a work in progress; the full representation of all values and variables needed for an arbitrary neuroimaging example will have to develop, though the initial project presented here addressed several complicated analyses. We have chosen at this time not to represent the distinction between image slices, image volumes, and the 4D imaging datasets which are the cornerstone of fMRI experiments, as they are not commonly included in analyses as variables per se; those structures will need to be developed and fleshed out to address the analyses at a more granular level.

The details of these structures and the related definitions are important to discuss within the fMRI ontology community. For example, we have represented the 'fMRI contrast image' variable as measuring the quality of 'a contrast between beta-values'. Another option would be more specific to the experiment in question, so that the contrast image measures 'the cortical activation increase with the oddball tones, based on the contrast between experimental conditions'. Both options are technically correct, and both are needed to compute properly over these representations of the experiment. It is crucial to the underlying representation, that these definitions are able to express what the contrast is about correctly and in context. It remains to be seen how these definitions relate to the cognitive paradigm and the mental processes allegedly under study. These are needed developments and should form the topic of discussion between specialists in the field and ontologists to ensure that the representations being built are relevant to the underlying scientific logic being used by scientists in their everyday work.

Several other ontological projects incorporate a representation of variables (or 'factors', or 'data elements') within their approach. The ISA-Tab (Investigation/Study/Assay) format uses a spreadsheet construct to capture and share data within a suite of tools (Rocca-Serra et al., 2010), but with no unifying ontology for the variables themselves. The 'Experimental Factor Ontology' (EFO) is a system to capture and track experimental factors within genomic experiments but has no consideration of measurement scales (Malone et al., 2010). The Ontology of Biomedical Investigations (OBI, previously mentioned) is an OBO-Foundry-authorized ontology (Brinkman et al., 2010), which may be considered a heavyweight complement to our lightweight methodology. We envisage that the curation processes we describe will make it easier for the formal ontologists working with OBI to collate terminology from researchers in the field, which they may then carefully organize into OBI. This then can form a concrete set of reference terms that we could then refer to in OoEVV.

OntoNeuroLog (Batrancourt et al., 2010) is a formal ontology, based on a different upper level ontology (DOLCE rather than BFO) with a different approach from the work we present here. Batrancourt et al. present a core ontology driven by the goal of sharing imaging data by providing formal structures for 'instruments' as an organizing principle. Instruments are all the materials and mechanisms by which we collect data in an experiment, and they tie very closely into the variables that they produce: An IQ-measuring instrument is a subject-data-acquisition instrument that produces an IQ measure as a variable. There are many points where OoEVV and OntoNeuroLog can interact; the distinction between numeric and non-numeric scores are represented in both frameworks, for example. We expect that as neuroimaging experiments are represented in OoEVV we should be able to submit new instruments and variables to various other ontologies for broader use, and to OntoNeuroLog for their use as well. The level of detailed involvement required to develop formal representations is both a necessary hindrance and a strength of detailed, formal approaches, however. With the OoEVV approach we finesse the differences between what is a subject-specific variable and what is a secondary variable, at the moment, for example, because we find it adds confusion. "Age" is a subject-specific variable when regressing brain volume against the subject's age, but it is an experimental factor when looking at the differences in measures on a new MRI scanner vs old—yet "age" refers to a similar quality in both cases. Even within a single experiment, "age" may be a primary variable in one analysis and a secondary variable or co-variate in another analysis; but that doesn't affect the representation of the experimental design. It plays a role in representing the statistical model for a given analysis, which is a different part of the larger framework for representing scientific knowledge. We anticipate that providing a simpler lightweight ODP as a way for neuroimagers to contribute their practical expertise will accelerate this larger effort towards increased formalization over time.

Within the OoEVV framework, we have divided classes of variables based on their 'measurement scales' (since this is needed to mathematically store and process variables' data). Another option would be to divide variables into subject-related measures such as clinical assessments and demographic measures, imaging-related measures such as the imaging coil used (when that is a covariate in the experimental analysis), and behavioral measures such as percent correct or reaction time, similar to what OntoNeuroLog does. This kind of categorization is useful when trying to list the variables collected in an experiment, to step through the aspects of the experiment and make sure no variable is missed. From the point of view of an analysis, however, those categories are less useful. Whether "age" is a demographic or clinical variable, for example, is not taken into account when an experimenter uses it in an ANOVA; however, whether age was recorded in years or categories such as "young", "older >40" or other groupings, is important both for analysis and for comparison across experiments. Thus we have focused on the scales, and what the scales measure, so that the quality of age can be measured by multiple scales.

To conclude, the development and involvement of a standardized representation for experimental variables, their values (and their involvement in experimental statistical findings) will be an important component for future neuroimaging data sharing architectures. The primary contributions of the solution we describe here are (a) a novel treatment of measurement scales; (b) efficient curation mechanism based on easy-to-use tools and (c) a centralized sharing mechanism through NCBO's BioPortal interface. We emphasize a lightweight approach that can be used in conjunction with more formal methods (such as OBI) and we provide a worked example for the fMRI imaging community.

Author contributions

Both authors contributed to this paper, with GAPCB developing OoEVV, KEfED and the overall framework, while JAT provided the domain examples identifying the neuroimaging needs and applications; both authors provide interactions with other ontologies' developers.

Acknowledgments

This work was supported by the BioScholar (GM083871) and the NeuARt (MH079068) projects (Dr. Burns, PI), and by a grant to Dr. Turner (1R01MH084812-01A1, NIMH). This work was also supported through the Biomedical Informatics Research Network (1 U24 RR025736-01, NCRR). BIRN provides a staged monitoring and classification process for biomedical informatics software development. The levels are named (a) 'Concept', (b) 'Prototype', (c) 'Limited availability' and (d) 'General availability', denoting the different stages of the support of a given capability by the BIRN technological approach. The OoEVV system is our instantiation of the capability to allow domain scientists to develop a local ontology of terminology within their current domain. Our current implementation of OoEVV is currently under evaluation for a level within this framework. We would also gratefully acknowledge feedback from Chris Stoeckert, Jie Zheng and Oliver He regarding the structure of OoEVV and its ability to interoperate with OBI.

Software availability and requirements
Project name: OoEVV
Project home page: http://www.isi.edu/projects/ooevv/overview
Operating system: MacOS, Windows, Linux
Programming language: Java
Other requirements: MySQL, Protégé
License: MIT

Conflict of interest

The authors state that they have no conflict of interest.

References

Andreasen, N., 1984. Scale for the Assessment of Positive Symptoms (SAPS). University of Iowa, Iowa City, IA.
Aranguren, M.E., Antezana, E., Kuiper, M., Stevens, R., 2008. Ontology design patterns for bio-ontologies: a case study on the cell cycle ontology. BMC Bioinformatics 9 (Suppl. 5), S1.
Batrancourt, B., Dojat, M., Gibaud, B., Kassel, G., 2010. A core ontology of instruments used for neurological, behavioral and cognitive assessments. Sixth International Conference on Formal Ontology in Information Systems (FOIS 2010), Presented at the formal Ontology in Information Systems (FOIS 2010), Toronto (Canada).
Blomqvist, E., 2009. Semi-automatic Ontology Construction Based on Patterns (Ph.D. thesis).
Brinkman, R.R., Courtot, M., Derom, D., Fostel, J.M., He, Y., Lord, P., Malone, J., Parkinson, H., Peters, B., Rocca-Serra, P., Ruttenberg, A., Sansone, S.A., Soldatova, L.N., Stoeckert, C.J.J., Turner, J.A., Zheng, J., 2010. Modeling biomedical experimental processes with OBI. J. Biomed. Semant. 1, S7.
Burns, G.A.P.C., Turner, J.A., 2012. Bottom-up curation of terminology for experimental variables: the Ontology of Experimental Variables and Values (OoEVV). Bio-Ontologies SIG, ISMB.
Ford, J.M., Roach, B.J., Jorgensen, K.W., Turner, J.A., Brown, G.G., Notestine, R., Bischoff-Grethe, A., Greve, D., Wible, C., Lauriello, J., Belger, A., Mueller, B.A., Calhoun, V., Preda, A., Keator, D., O'Leary, D.S., Lim, K.O., Glover, G., Potkin, S.G., Mathalon, D.H., 2009. Tuning in to the voices: a multisite FMRI study of auditory hallucinations. Schizophr. Bull. 35, 58–66.
Fox, P.T., Laird, A.R., Fox, S.P., Fox, P.M., Uecker, A.M., Crank, M., Koenig, S.F., Lancaster, J.L., 2005. BrainMap taxonomy of experimental design: description and evaluation. Hum. Brain Mapp. 25, 185–198.
Gadde, S., Aucoin, N., Grethe, J.S., Keator, D.B., Marcus, D.S., Pieper, S., 2012. XCEDE: an extensible schema for biomedical data. Neuroinformatics 10, 19–32.
Gangemi, A., Presutti, V., 2009. Ontology design patterns. In: Staab, S., Studer, R. (Eds.), Handbook of Ontologies. Springer.
Glahn, D.C., Ragland, J.D., Abramoff, A., Barrett, J., Laird, A.R., Bearden, C.E., Velligan, D.I., 2005. Beyond hypofrontality: a quantitative meta-analysis of functional neuroimaging studies of working memory in schizophrenia. Hum. Brain Mapp. 25, 60–69.
Groth, P., Gibson, A., Velterop, J., 2010. The anatomy of a nanopublication. Inform. Serv. Use 30, 51–56.
Helmer, K., Ambite, J., Ames, J., Ananthakrishnan, R., Burns, G., Chervenak, A., Foster, I., Liming, L., Keator, D., Macciardi, F., Madduri, R., Navarro, J., Potkin, S., Rosen, B., Ruffins, S., Schuler, R., Turner, J., Toga, A., Williams, C., Kesselman, C., 2011. Enabling collaborative research using the Biomedical Informatics Research Network (BIRN). J. Am. Med. Inform. Assoc. 18, 416.
Hollingshead, A., Redlich, F., 1958. Social Class and Mental Illness: A Community Sample. John Wiley & Sons, Ltd., New York.
Laird, A.R., Lancaster, J.L., Fox, P.T., 2005. BrainMap: the social evolution of a human brain mapping database. Neuroinformatics 3, 65–78.
Laird, A.R., Eickhoff, S.B., Kurth, F., Fox, P.M., Uecker, A.M., Turner, J.A., Robinson, J.L., Lancaster, J.L., Fox, P.T., 2009. ALE meta-analysis workflows via the BrainMap database: progress towards a probabilistic functional brain atlas. Front. Neuroinform. 3, 23.
Laird, A.R., Eickhoff, S.B., Fox, P.M., Uecker, A.M., Ray, K.L., Saenz Jr., J.J., McKay, D.R., Bzdok, D., Laird, R.W., Robinson, J.L., Turner, J.A., Turkeltaub, P.E., Lancaster, J.L., Fox, P.T., 2011. The BrainMap strategy for standardization, sharing, and meta-analysis of neuroimaging data. BMC Res. Notes 4, 349.
Malone, J., Holloway, E., Adamusiak, T., Kapushesky, M., Zheng, J., Kolesnikov, N., Zhukova, A., Brazma, A., Parkinson, H., 2010. Modeling sample variables with an experimental factor ontology. Bioinformatics 26, 1112–1118.
Michael, B., Robert, L., Gibbon, M., Williams, J.B.W., 2002. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P). Biometrics Research, New York State Psychiatric Institute, New York.
Minzenberg, M.J., Laird, A.R., Thelen, S., Carter, C.S., Glahn, D.C., 2009. Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. Arch. Gen. Psychiatry 66, 811–822.
Mungall, C.J., Gkoutos, G.V., Smith, C.L., Haendel, M.A., Lewis, S.E., Ashburner, M., 2010. Integrating phenotype ontologies across multiple species. Genome Biol. 11, R2.
Poldrack, R.A., Fletcher, P.C., Henson, R.N., Worsley, K.J., Brett, M., Nichols, T.E., 2008. Guidelines for reporting an fMRI study. Neuroimage 40, 409–414.
Rocca-Serra, P., Brandizi, M., Maguire, E., Sklyar, N., Taylor, C., Begley, K., Field, D., Harris, S., Hide, W., Hofmann, O., Neumann, S., Sterk, P., Tong, W., Sansone, S.A., 2010. ISA software suite: supporting standards-compliant experimental annotation and enabling curation at the community level. Bioinformatics 26, 2354–2356.
Russ, T., Ramakrishnan, C., Hovy, E., Bota, M., Burns, G., 2011. Knowledge engineering tools for reasoning with scientific observations and interpretations: a neural connectivity use case. BMC Bioinformatics 12, 351.
Saver, J.L., Warach, S., Janis, S., Odenkirchen, J., Becker, K., Benavente, O., Broderick, J., Dromerick, A.W., Duncan, P., Elkind, M.S.V., Johnston, K., Kidwell, C.S., Meschia, J.F., Schwamm, L., 2012. Standardizing the structure of stroke clinical and epidemiologic research data: the National Institute of Neurological Disorders and Stroke (NINDS) Stroke Common Data Element (CDE) project. Stroke 43 (4), 967–973.
Stevens, S.S., 1946. On the theory of scales of measurement. Science 103, 677–680.
Tallis, M., Dave, D., Burns, G.A., 2012. Preliminary meta-analyses of experimental design with examples from HIV vaccine protection studies. Discovery Informatics Symposium DIS2012. Presented at the Discovery Informatics Symposium DIS2012, Arlington VA.
Turner, J.A., Laird, A.R., 2012. The cognitive paradigm ontology: design and application. Neuroinformatics 10, 57–66.
Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., Wager, T.D., 2011. Large-scale automated synthesis of human functional neuroimaging data. Nat. Methods 8, 665–670.